AD_____

Award Number: DAMD17-98-1-8513

TITLE: Exercise and Bone Density: Meta-Analysis

PRINCIPAL INVESTIGATOR: George A. Kelley, M.D.

CONTRACTING ORGANIZATION: Massachusetts General Hospital
Boston, MA 02114-2554

REPORT DATE: October 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2003	Annual (1 Sep 2002 - 1 Sep 2003)

4. TITLE AND SUBTITLE

Exercise and Bone Density: Meta-Analysis

5. FUNDING NUMBERS

DAMD17-98-1-8513

6. AUTHOR(S)

George A. Kelley, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Massachusetts General Hospital
Boston, MA 02114-2554

E-Mail: gkelley@hsc.wvu.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Purpose: Use the meta-analytic approach to examine the effects of exercise on bone mineral density (BMD) using individual patient data (IPD) and to compare these results with summary data from the same studies. **Scope:** The specific aim for this project period was to analyze IPD from existing exercise training studies dealing with the effects of exercise on BMD. **Major Findings:** None for this reporting period. We are still awaiting transfer of this grant from Massachusetts General Hospital, Boston MA to West Virginia University, Morgantown, WV. Consequently, we have been unable to work on this project for the current reporting period. We plan on completing this project when the grant is transferred to West Virginia University. Details are provided in our cover letter.

14. SUBJECT TERMS

Exercise, bone, meta-analysis, osteoporosis, bone density, physical activity, physical fitness, men, women

15. NUMBER OF PAGES

37

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

I. INTRODUCTION

Bone health is critical for optimal performance and the prevention of fractures associated with low bone mineral density (BMD). Our first two-year period of funding focused on using the meta-analytic approach to examine the effects of exercise on BMD in adult humans using summary means from completed studies. Since no meta-analysis had existed using individual patient data (IPD) to examine the effects of exercise on BMD, our second period of funding was devoted to examining the feasibility of such. The specific aims of the second project period were to (1) compare summary versus IPD in relation to the overall magnitude of effect that exercise has on BMD, (2) compare summary versus IPD in relation to the effect of potentially confounding variables (age, training program, etc.) on changes in BMD, and (3) provide recommendations for future research regarding the use of summary versus IPD for examining the effects of exercise on BMD. The results of this project will help identify the best approach to use (summary versus IPD) when attempting to arrive at a more objective conclusion regarding the effects of exercise on BMD in humans. In addition, this will be the first meta-analysis using IPD in the area of exercise and BMD. Finally, the results of this project will provide the Armed Forces with a better understanding of the effects of exercise on BMD and will also help to identify what programs, if any, will provide for optimum bone development and maintenance.

II. BODY

A. Statement of Work

None. No work was conducted for the past year because this grant has not yet been transferred from Massachusetts General Hospital in Boston to West Virginia University in Morgantown.

III. KEY RESEARCH ACCOMPLISHMENTS

None. No work was conducted for the past year because this grant has not yet been transferred from Massachusetts General Hospital in Boston to West Virginia University in Morgantown.

IV. REPORTABLE OUTCOMES (Note: The items referenced below represent work which was conducted prior to the current reporting year and previously listed as "in press")

A. Manuscripts (Refereed)

1. Kelley, G.A., Kelley, K.S., Tran, Z.V. Aerobic exercise and regional bone density in women: A meta-analysis of controlled trials. American Journal of Medicine & Sports. 2002; 4:427-433. (See Appendix A)

2. Kelley, G.A., Kelley, K.S., Tran, Z.V. Retrieval of individual patient data for an exercise-related meta-analysis. American Journal of Medicine and Sports. 2002; 4:350-354. (See Appendix B)

3. Kelley, G.A., Kelley, K.S., Tran, Z.V. Exercise and lumbar spine bone mineral density in postmenopausal women: A meta-analysis of individual patient data. Journal of Gerontology: Medical Sciences. 2002; 57A:M599-M604. (See Appendix C)

4. Kelley, G.A., Kelley, K.S., Tran, Z.V. Exercise and bone mineral density at the femoral neck in postmenopausal women: A meta-analysis of controlled clinical trials using individual patient data. Disease Management and Clinical Outcomes. **Note:** This manuscript, previously reported as "in press," will not be published in the above-referenced journal because it has ceased publication. We plan on seeking publication of this work elsewhere. For verification regarding this journal, please contact Dr. Clifton Lacy (Phone: 609-292-7837; E-mail: lacy1@optonline.net) (See Appendix D)

V. CONCLUSIONS

A. Importance of Completed Research

No work was conducted for the past year because this grant has not yet been transferred from Massachusetts General Hospital in Boston to West Virginia University in Morgantown.

B. Suggestions for Future Work

A need exists for addressing the inconsistencies between results of summary means and individual patient data in relation to the effects of exercise at the hip. This includes an examination of whether the smaller number of individual patient data studies available at the femoral neck limited the power of our more recent meta-analysis or the inclusion of studies using any femur site skewed the results of our earlier meta-analytic work.

C. So What?

This cannot be addressed because no work was conducted for the past year on this project. We are still awaiting transfer of this grant from Massachusetts General Hospital in Boston to West Virginia University in Morgantown.

VI. REFERENCES – Not Applicable

VII. APPENDICES

A. Aerobic Exercise and Regional Bone Density Manuscript

B. Retrieval of Individual Patient Data Manuscript

C. Exercise and Lumbar Spine Bone Mineral Density Manuscript

D. Exercise and Bone Mineral Density at the Femoral Neck Manuscript

APPENDIX A

Aerobic Exercise and Regional Bone Density in Women: A Meta-Analysis of Controlled Trials

George A. Kelley, DA; Kristi S. Kelley, MEd
Institute of Health Professions, Massachusetts General Hospital, Boston, MA

In this study the meta-analytic approach was used to examine the effects of aerobic exercise on regional bone mineral density at the lumbar spine, femur, and radius in women. Twenty-four studies representing 58 groups (31 exercise, 27 control) and 1029 subjects (517 exercise, 512 control) met the criteria for inclusion. Using a random-effects model, small but statistically significant effect size changes in bone mineral density were observed at the lumbar spine ($\bar{X}$ ±SD=0.33±0.49; 95% confidence interval=0.16–0.50) and femur ($\bar{X}$ ±SD=0.25±0.35; 95% confidence interval=0.14–0.35). Changes in lumbar spine bone mineral density were equivalent to a 0.37% increase in the exercise groups and a 1.87% decrease in the control groups. For the femur, changes were equivalent to a 1.37% increase in the exercise groups and a 0.58% decrease in the control groups. No statistically significant changes were observed at the radius ($\bar{X}$ ±SD=0.10±0.45; 95% confidence interval=−0.20 to 0.41). The overall results of this study suggest that aerobic exercise has a small but positive effect on bone mineral density at the lumbar spine and femur in women. (Am J Med Sports. 2002; 4:427–433, 452)

Address for correspondence:
George A. Kelley, DA, Associate Professor, Graduate Program in Clinical Investigation; Director, Meta-Analytic Research Group, Institute of Health Professions, Massachusetts General Hospital, 101 Merrimac Street, Boston, MA 02214
Manuscript received August 15, 2000; accepted August 30, 2000

Osteoporosis, defined as abnormally low bone mass, is a major public health problem in the United States, as well as other countries. In 1996, it was estimated that approximately 23 million women in the United States had osteoporosis or were at risk for developing the disease.[1] By the year 2015 this figure is expected to increase to approximately 35 million.[2] It is well established that low bone mineral density (BMD) is associated with increased fracture risk. The health care costs associated with osteoporotic fractures has been reported to exceed $13.8 billion annually.[3] Given the health and economic costs associated with osteoporosis, a need exists for appropriate nonpharmacologic and pharmacologic interventions for dealing with this disease. One such nonpharmacologic intervention may be aerobic exercise,[4] a low-cost intervention that is available to most of the general public.

We have previously reported that aerobic exercise might help to maintain and/or increase BMD in postmenopausal women but that additional studies were needed before any firm conclusions could be reached.[5–7] Since the time of these published meta-analyses, a number of additional studies have been conducted and/or located. It is critical that up-to-date meta-analyses be performed in order to provide the most recent information possible on the state of knowledge regarding the topic of interest. Given the health care consequences of low BMD, it is important to understand the role that aerobic exercise may play as a nonpharmacologic intervention for enhancing and/or maintaining BMD in women. Thus, we used the meta-analytic approach to examine the effects of aerobic exercise on regional BMD at the lumbar spine, femur, and radius in women.

Methods

DATA SOURCES. Computerized literature searches of articles indexed between January, 1966 and December, 1998 were performed using MEDLINE, Embase, Current Contents, Sport Discus, and Dissertation Abstracts International Databases. The key words used in this literature search were "exercise" and "bone."

While this broad approach to searching the literature will result in the retrieval of a greater number of articles to review, it should decrease the number of studies missed when a more narrow and focused search is conducted. In addition to computerized literature searches, the reference lists from both original and review articles were examined in order to identify any studies that had not been previously identified and that appeared to contain information that may have met our inclusion criteria. Finally, three experts on exercise and bone density (Dr. David Nichols, Dr. Charlotte Sanborn, and Dr. Christine Snow) reviewed our reference list for thoroughness and completeness.

STUDY SELECTION. The inclusion criteria for this study were as follows: 1) trials were randomized or nonrandomized trials and included a comparative nonexercise group; 2) aerobic exercise was the only intervention; 3) subjects were adult female humans (mean age, 18 years or older); 4) studies were reported as journal articles, dissertations, and master's theses published in the English language literature; 5) studies were published and indexed between January, 1966 and December, 1998; 6) BMD (relative value of bone mineral per measured bone area) was assessed at the femur, lumbar spine, or radius; and 7) training studies which lasted a minimum of 16 weeks. Only studies that met the above criteria were included in our analysis. Thus, for example, if BMD was also assessed in women performing progressive resistance exercise as the primary training modality, we did not include this information since it did not meet our inclusion criteria. Because dissertations and master's theses may eventually become full-length journal articles, we cross-referenced between the two in order to avoid duplication. We did not include abstracts and conference papers from national meetings because of the paucity of data provided as well as the inability to obtain complete data from the authors. Studies published in foreign language journals were also not included because of the potential error in the translation and interpretation of findings. Studies that met our inclusion criteria were also examined to ensure that the same subjects were not included in more than one study.[8] For studies that met our inclusion criteria but did not provide appropriate information on changes in BMD, personal contact was made with the authors in an attempt to retrieve such information.

DATA ABSTRACTION. Coding sheets that could hold 242 items per study were developed and utilized in this study. In order to avoid inter-coder bias, all data were independently abstracted by both authors. The authors then met and reviewed every data point for accuracy and consistency. Disagreements were resolved by consensus. The major categories of variables coded included study characteristics, physical characteristics of subjects, and primary and secondary outcomes.

STATISTICAL ANALYSIS. ***Primary Outcomes.*** The primary outcomes in this study were changes in BMD at the lumbar spine, femur, and radius, calculated using the standardized difference effect size (ES) approach. This was accomplished by subtracting the change outcome in the exercise group from the change outcome in the control group, then dividing this difference by the pooled standard deviation of the exercise and control groups.[9] This measure provides one with a statistic similar to a *z* score. In general, an ES of 0.20 is considered a small effect, 0.50 a moderate effect, and 0.80 a large effect.[10] An ES of 0.30 for example, means that the exercise group differed from the control group by three-tenths of a standard deviation in favor of the exercise group. Using a *z* score table, this means that the exercise group would do better than approximately 62% of the control group. We used this approach vs. the original metric because of the various ways in which the authors reported data on changes in BMD and because we also wanted to maximize the number of studies and outcomes that could be included in our analysis. All ESs were then corrected for small-sample bias.[9] For studies that did not report change outcome variances, these were estimated using previously developed methods.[11] T-distribution 95% confidence intervals (CI) were calculated for all outcomes. If the 95% CI included zero (0.00), it was concluded that there was no statistically significant effect of exercise on BMD. A random-effects model was used for all analyses.[9]

Heterogeneity of ESs was examined using the Q statistic.[9] For studies that included multiple outcomes because of more than one group, net changes were initially treated as independent data points. However, in order to examine the influence (sensitivity) of each study on the overall results, analyses were performed with each study deleted from the model for ES changes at the lumbar spine, femur, and radius. Publication bias (the tendency for journals and/or authors to publish studies that yield statistically significant results) was examined using a funnel plot.[12] This was accomplished by plotting the sample size on the vertical axis and ES changes in BMD on the horizontal axis. Usually, smaller studies tend to disperse at the bottom of the funnel while larger studies tend to congregate at the top. A gap at the bottom of the funnel on the left side indicates that small studies yielding null or negative results may be missing. Study quality was assessed using a three-item questionnaire designed to assess bias—specifically, randomization, blinding, and withdrawals/dropouts.[13] The number of points possible ranged from a low of 0 to a high of 5. All

questions were designed to elicit yes (1 point) or no (0 points) responses. The questionnaire, which took less than 10 minutes per study, has been shown to be both valid (face validity) and reliable (researcher inter-rater agreement: r=0.77; 95% CI=0.60–0.86).[13]

Subgroup Analyses. Subgroup analyses for ES changes at the lumbar spine and femur were performed using analysis of variance (ANOVA)-like procedures for meta-analysis.[9] These procedures provide statistics for both within (Q_w)- and between (Q_b)-group differences. A random-effects model was used for all analyses. Subgroup analyses were performed for: ES changes at the lumbar spine and femur according to type of publication (journal vs. dissertation); country in which the study was conducted (United States vs. other); study design (randomized vs. nonrandomized, controlled trial); whether subjects were postmenopausal; whether subjects were taking calcium supplementation; type of BMD assessment (dual-energy x-ray absorptiometry, dual photon absorptiometry, quantitative computed tomography); and higher vs. lower impact activity. Higher impact activities included exercises such as running, jumping, and aerobic dancing with both feet off the ground, while lower impact activities included exercises such as walking and low impact aerobic dancing with both feet on the ground. ES changes in BMD at the femur were also examined when data were partitioned according to whether drugs were taken that could enhance BMD, cigarette smoking, diet, previous physical activity, and the specific site of BMD assessment (femoral neck, trochanter, Ward's triangle, intertrochanter). Insufficient data were provided to examine ES changes in BMD at the lumbar spine according to whether drugs were taken that could enhance BMD, cigarette smoking, diet, previous physical activity, and the specific site at which BMD was assessed. For both the lumbar spine and femur, insufficient data were provided to examine changes in BMD when partitioned according to alcohol consumption and previous fractures. We were unable to partition the results according to training modality because of the variety of activities in which the subjects participated. We did not perform subgroup analysis for changes in BMD at the radius because of the small sample size. In addition, we were not able to examine differences between the radius and other sites at the forearm (for example, the ulna) because of insufficient data.

Regression Analysis. The potential associations between ES changes in BMD at the lumber spine and femur were conducted using simple weighted least-squares regression, according to procedures developed by Hedges and Olkin.[9] Variables included study quality, percent dropout, initial BMD, age, height, initial body weight, changes in body weight, initial body mass index, changes in body mass index, initial percent fat, changes in percent fat, initial lean body mass, changes in lean body mass, initial maximum oxygen consumption ($mL/kg^{-1}/min^{-1}$), changes in maximum oxygen consumption ($mL/kg^{-1}/min^{-1}$), years past menopause, initial calcium intake, changes in calcium intake, reliability of BMD measurements, length, frequency, intensity, and duration of training, total minutes of training (length × frequency × duration), and compliance, defined as the percentage of exercise sessions attended. Insufficient data were available to examine ES changes in BMD and resting heart rate. We did not conduct regression analyses for ES changes in BMD at the radius because of the small sample size. We were unable to conduct any type of multiple regression analyses because of missing data for different sets of variables.

Secondary Outcomes. Secondary outcomes (changes in body weight, body mass index, percent body fat, lean body mass, maximum oxygen consumption, resting heart rate, and calcium intake) were calculated as the difference (exercise minus control) of the changes (initial minus final) in these mean values. With the exception of the use of the original metric vs. standardized difference approach, changes in secondary outcomes were examined using the same procedures as those for BMD.

An independent t test (2-tailed) was used to compare differences in study quality between journals and dissertations. Unless otherwise noted, all results are reported as $\bar{X} \pm SD$. The alpha level for statistical significance was set at $p \leq 0.05$.

Results

STUDY CHARACTERISTICS. Twenty-seven studies met the criteria for inclusion[14–40]; however, we were unable to retrieve necessary data from three studies.[15,27,28] This resulted in a loss of approximately 11%. Thus, 24 studies representing 31 exercise and 27 control groups (some studies had more than one group) were included in our final analysis.[14,16–26,29–40] From these 24 studies, 31 effect sizes were generated for the lumbar spine, 42 for the femur, and 11 for the radius. Twenty-two of the studies were published in refereed journals[14,17–26,29,31–40] while the other two were dissertations.[16,30] Thirteen studies were conducted in the United States,[14,16,19,20,22,26,30,31,33,34,37–39] three each in Australia[18,35,36] and the United Kingdom,[17,21,32] two each in Finland[24,25] and Japan,[23,40] and one in China.[29] Thirteen of the studies were randomized, controlled trials,[16,17,21–25,29,31,32,35,36,39] while 11 were nonrandomized, controlled trials.[14,18–20,26,30,33,34,37,38,40] Study quality ranged from 0

to 5 ($\bar{X}$±SD=1.75±1.51). There was no statistically significant difference in study quality between studies published in journals and dissertations (p=0.65). A total of 1029 subjects (517 exercise, 512 control) completed pre- and post-assessments of BMD. The average number of subjects ranged from five to 49 in the exercise groups ($\bar{X}$±SD=17±12) and from four to 48 in the control groups ($\bar{X}$±SD=19±15). The percent dropout, defined as the percentage of subjects who did not complete the study, ranged from 0%–63% in the exercise groups ($\bar{X}$±SD=20%±16%) and from 0%–43% in the control groups ($\bar{X}$±SD=10%±11%).

SUBJECT CHARACTERISTICS. A description of the subject characteristics is shown in Table I. In six studies, all of the subjects were white[14,22,30,31,33,36]; in one study, all subjects with the exception of one (a black person) were white[20]; in one study, all subjects were Chinese[29]; and in one study, all subjects were Japanese.[40] In 19 studies, all subjects were postmenopausal[14,17–23,26,29–33,35–38,40]; in two studies, only some subjects were postmenopausal[25,34]; and in three studies, no subjects were postmenopausal.[16,24,39] In 14 studies, no subjects were taking any type of hormone replacement during the study[14,17–19,23, 30–36,38,40] and in six studies, some of the subjects were taking some type of hormone replacement therapy.[20–22,24,25,37] One study had two separate groups of subjects in which one group took some type hormone replacement therapy while the other did not.[26] In nine studies, all subjects were taking some type of calcium supplementation during the study[16,20,26,31–33,35,37,39]; in five studies, no subjects were taking any type of calcium supplementation[19,22,23,36,40]; and in two studies, some of the subjects took some type of calcium supplementation.[21,30] Another study had two separate groups of subjects, one who took some type of calcium supplementation and another who did not.[29] In one other study, all of the subjects in the control group took some type of calcium supplementation, while some in the exercise group did so.[14] In eight studies, food intake did not change during the study[14,17,19,22,24,26,32,38] and in one study, it did.[33] In six studies, none of the subjects smoked cigarettes[24–26,30,36,38] and in four studies, some of the subjects smoked.[17,19,21,37] In one study, none of the subjects in the control group smoked but some of the subjects in one of the two exercise groups smoked.[31] In another study, some of the subjects in the exercise group smoked but no subjects in the control group did.[18] In two studies, some subjects consumed alcohol during the study.[18,32] In two studies, no subjects had previous fractures,[29,38] while in another study, subjects did have previous fractures.[21] In 13 studies, none of the subjects had been active prior to taking part in the study[14,17,20, 22,24–26,30,31,33,37,38,40] and in six studies, some of the subjects had been previously active.[16,19,21,35,36,39] In one study, no subjects in the control group had been active prior to taking part in the study but subjects in the exercise group had been previously active.[34]

Table I. Subject Characteristics

Variable	N	Exercise ($\bar{X}$±SD)	N	Control ($\bar{X}$±SD)
Age (years)	31	57.9±12.7	27	58.2±13.2
Height (cm)	22	160.7±4.3	19	161.5±4.4
Weight (kg)	25	64.7±6.6	21	64.2±6.4
BMI (kg/m^2)	24	24.9±1.9	21	24.6±1.9
Fat (%)	13	38.2±4.8	10	37.9±6.5
Lean mass (kg)	13	41.2±3.5	10	39.8±2.8
Initial $V_{O_{2max}}$ (mL/kg^{-1}min^{-1})	16	23.4±4.2	11	23.9±5.0
Initial RHR (bpm)	4	76.7±3.7	2	74.15±4.5
Postmenopausal (years)	22	10.0±5.4	18	11.7±5.8
Calcium (mg)	19	934±340	16	938±344

N=number of groups reporting mean data; BMI=body mass index; $V_{O_{2max}}$=maximal oxygen consumption; RHR=resting heart rate; bpm=beats per minute.

BONE DENSITY ASSESSMENT CHARACTERISTICS. Twelve studies assessed BMD at the lumbar spine using dual-energy x-ray absorptiometry (DEXA),[16,17,21,23–26,29,35,38–40] seven studies used dual-photon absorptiometry (DPA),[14,18,20,22,30,31,34] and two used quantitative computed tomography (QCT).[19,32] One other study used both DPA and QCT to assess BMD at the lumber spine.[33] For studies that included such data, the vast majority reported the assessment of BMD at the L2–L4 sites.[16,17,20,22–25,29–31,34,38,39] Three studies reported the assessment of BMD at the L1–L4 sites,[14,35,40] one at the L1–L2 sites,[19] and another at the L1–L3 and L2–L4 sites.[33] Between-study mean reliability (coefficient of variation) of BMD assessment at the lumber spine ranged from 0.4%–3%. Ten studies used DEXA to assess BMD at the femur,[16,17,21,24–26,29,35,38,39] while another five used DPA.[14,18,30,33,34] Fifteen studies included assessment of BMD at the femoral neck,[14,16–18,21,24–26,29,30,33–35,38,39] seven at Ward's triangle,[16,18,26,29,34,38,39] eight at the trochanter,[16,18,24,26,34,35,38,39] and two at the intertrochanter.[29,35] One study involved BMD assessment at the distal femur,[24] and another involved assessment of the total femur.[35] The mean between-study reliability (coefficient of variation) for BMD assessment at the femur ranged from 0.5%–4.4%. In eight studies, BMD was assessed at the forearm[24–26,30,33,34,36,37]; however, we were unable to identify whether one of the studies assessed BMD at the radius.[36] Four studies used single-photon

absorptiometry (SPA) to assess BMD at the radius,[30,33,34,37] while three used DEXA.[24–26] The mean between-study reliability (coefficient of variation) ranged from 0.5%–5.0%.

TRAINING PROGRAM CHARACTERISTICS. A description of the training program characteristics is shown in Table II. Overall, the activity most commonly included in these exercise interventions was walking. Specifically, five studies limited the training modality to primarily walking,[17,19,21,23,33] two to jogging,[16,39] and two to a combination of walking and jogging.[31,38] Two other studies had subjects participate primarily in aerobic dancing,[32,34] while another two employed walking[35,36] or aerobic dancing[18,22] as well as other activities. One study limited participants' exercise to stair stepping and other miscellaneous activities,[29] while another limited exercise to stationary cycling.[14] Two other studies had participants take part in a combination of walking, jogging, cycling, stair stepping, and other activities[20,25]; one had subjects perform walking, jogging, and stair stepping[26]; and another had subjects walk, swim, and perform other various activities.[40] One study had subjects perform aerobic dancing, stair stepping, and other assorted activities,[24] while another had subjects perform a variety of different but unspecified activities.[37] Finally, one study had one group of subjects who walked and another group who swam.[30]

PRIMARY OUTCOMES. *Lumbar Spine.* The overall results for ES changes in lumbar spine BMD are shown in Table III. As can be seen, small but statistically significant ES changes in lumber spine BMD were observed. These changes were equivalent to a 0.37% increase in the exercise groups and a 1.87% decrease in the control groups. No statistically significant heterogeneity was found for changes in lumbar spine BMD. Funnel plot analysis was suggestive of publication bias. With each study deleted from the model once, ES changes in BMD ranged from a low of 0.27±0.42 (95% CI=0.12–0.44) to a high of 0.36±0.48 (95% CI=0.18–0.54).

Femur. The overall results for ES changes in BMD at the femur are shown in Table III. As can be seen, small but statistically significant changes in BMD at the femur were observed. These changes were equivalent to a 1.37% increase in the exercise groups and a 0.58% decrease in the control groups. No statistically significant heterogeneity was found for changes in BMD at the femur. Funnel plot analysis was suggestive of publication bias. With each study deleted from the model once, ES changes in BMD at the femur ranged from a low of 0.21±0.34 (95% CI=0.10–0.32) to a high of 0.26±0.38 (95% CI=0.14–0.38).

Table II. Training Program Characteristics

VARIABLE	N	($\bar{X}$ ±SD)
Length (weeks)	31	53±23
Frequency (days/week)	28	3±1
Intensity (% Vo_{2max})	7	75±8
Duration (min/session)	22	33±11
Total min*	22	5046±3159
Compliance (%)	21	83±12

N=number of groups reporting mean data; Vo_{2max}=maximal oxygen consumption; *total minutes calculated as the product of length, frequency, and duration.

Radius. The overall results for ES changes in BMD at the radius are shown in Table III. As can be seen, changes in BMD at the radius were not statistically significant. ES changes were equivalent to a 0.08% decrease in BMD for the exercise groups and a 0.75% decrease in the control groups. No statistically significant heterogeneity was found for changes in BMD at the radius. Funnel plot analysis was not suggestive of publication bias. With each study deleted from the model once, ES changes in BMD at the radius ranged from a low of 0.02±0.37 (95% CI=−0.25 to 0.28) to a high of 0.17±0.42 (95% CI=−0.13 to 0.48).

Subgroup and Regression Analysis. Greater ES changes in BMD at the femur were observed for those subjects who received some type of calcium supplementation ($\bar{X}$ ±SD, calcium supplementation=0.33±0.42; no calcium supplementation, −0.24±0.44; Q_b=4.55; p=0.03). None of the other subgroup analyses at the lumber spine and femur were statistically significant or clinically important.

SECONDARY OUTCOMES. A statistically significant increase was observed for changes in maximum oxygen consumption ($\bar{X}$ ±SD=1.86±2.17 mL/kg^{-1}min^{-1}; 95% CI=0.31–3.41). No statistically significant or clinically important changes were found for any of the other secondary outcomes.

Discussion

One of the primary roles of a meta-analysis is to attempt to arrive at some overall conclusion(s) regarding a particular body of research. The overall results of this study suggest that aerobic exercise has a small but positive effect on BMD at the lumbar spine and femur in both premenopausal and postmenopausal women, and that this effect appears to be the result of increasing and/or preserving BMD. The fact that a similar effect was not found at the radius is not surprising, given that it appeared that all of the exercise interventions that the studies employed focused on

Table III. Overall Results for BMD

VARIABLE	ES(#)	$\bar{X}$±SD	95% CI	Q(P)
Lumbar spine	31	0.33±0.49	0.16 to 0.50*	33.65 (0.29)
Femur	42	0.25±0.35	0.14 to 0.35*	32.93 (0.81)
Radius	10	0.10±0.45	−0.20 to 0.41	09.99 (0.44)

BMD=bone mineral density; CI=confidence interval.
* Significantly different from zero.

loading the lower extremities. Thus, specific loading at all sites, including the radius, may be necessary in order to help increase and/or preserve BMD at that particular site. The overall results observed in this study are similar to those of our previous and less complete work, in which comparable changes in BMD were reported.[5–7]

While the results of this study are positive with respect to changes in BMD at the lumbar spine and femur, the clinical importance of such small changes (approximately 2%) is not known, especially as they relate to fracture risk. Indeed, it may be that postmenopausal women might need other types of nonpharmacologic and pharmacologic interventions in addition to, or in lieu of, aerobic exercise in order to realize a significant impact on increasing and/or preserving BMD and subsequently reducing fracture risk. For example, a recent meta-analysis found that 10 mg per day of alendronate over a period of 3 years in postmenopausal, osteoporotic women reduced the estimated cumulative incidence of nonvertebral fractures from 12.6% in the placebo group to 9.0% in the alendronate group.[41] This coincided with an increase in BMD of approximately 8.8% at the spine, 7.8% at the trochanter, and 5.9% at the femoral neck.[41] Since the changes in BMD observed in this meta-analysis were much smaller, it is difficult to generalize as to how these changes impact subsequent fracture risk. It would appear plausible to suggest that future studies examining the effects of exercise on changes in BMD attempt to address the clinical importance of these changes on subsequent fracture risk.

The fact that there were greater changes in BMD at the femur in those studies that included calcium supplementation suggests that its combination with exercise may be necessary in order to increase and/or preserve BMD in women. This supports previous work in which calcium supplementation was found to be necessary in order to maximize the benefits of exercise on BMD.[42]

We were surprised to find that both higher and lower impact activity yielded similar benefits at both the femur and lumbar spine, especially since it is generally believed that higher impact activity will have a more positive effect on BMD. However, our results support other reports of similar BMD results for both higher and lower impact activities.[22] This notwithstanding, our results need to be interpreted with caution, since the issue of mechanical loading and skeletal integrity is still a controversial area in need of additional research.[43] Furthermore, since few authors reported the specific ground-reaction forces associated with the intervention employed, we were limited to developing a somewhat arbitrary classification system.

Despite the fact that meta-analysis is a quantitative approach for reviewing a body of literature, subjective decisions still have to be made. For example, in this investigation, we chose to include unpublished studies (dissertations) in our analysis. While the inclusion of unpublished studies in scientific overviews is controversial, we believe that if appropriate resources are available, unpublished studies should not be systematically excluded. Rather, they should be included and examined for potential differences when compared to published work. This is especially true given the fact that there is a bias toward publishing studies that yield statistically significant and positive results. For example, Sterling et al.[44] found that approximately 96% of selected psychology journals and 85% of selected medical journals published studies that yielded a statistically significant result. The inclusion of unpublished work in scientific overviews is a feeling that is shared by the vast majority of meta-analysts and methodologists; approximately 78% believe that unpublished material should definitely or probably be included in scientific overviews.[45] Alternatively, it may be argued that the inclusion of unpublished work is inappropriate because it has not gone through the peer review process and/or that such studies were never submitted for publication consideration because of the feeling that they may have been flawed because of some type of methodologic problem. However, the fact that we found no statistically significant difference in study quality between published and unpublished work and found no difference in ES results when our data were partitioned according to type of publication led us to include this information in our analysis.

Another subjective decision we made was the inclusion of nonrandomized, controlled trials. We believe that it is important to include nonrandomized trials, at least in the exploratory phase, in order to see if they differ from randomized trials. Since our subgroup analyses revealed no statistically significant differences in ES between randomized and nonrandomized trials at any of the sites assessed, we chose to include these in our final analysis.

While it appears that aerobic, site-specific exercise has a small but positive effect on BMD in adult women, these results need to be interpreted with regard to the following caveats. First, the fact that our

funnel plot analysis was suggestive of publication bias for both lumbar spine and femur results may warrant caution in the interpretation of our findings. We chose to use this quasi-statistical approach because the statistical approaches that have been developed to date are not grounded in formal statistical theory and make assumptions that are doubtful or indefensible.[46] However, it is also important to realize that the sensitivity of funnel plots for detecting publication bias has not been assessed systematically.[46] Second, the very nature of meta-analysis dictates that the meta-analysis itself inherits the limitations of the studies included in the analysis. For example, we were unable to perform subgroup analyses of ES changes in BMD at the lumbar spine according to whether drugs were taken that could enhance BMD, cigarette smoking, diet, previous physical activity habits, and the specific site at which BMD was assessed. In addition, insufficient information was available to examine ES changes in BMD at both the lumbar spine and femur according to alcohol consumption, previous fractures, and training modality. Furthermore, we were limited to conducting simple vs. multiple regression analysis because of missing data. The ability to include this missing information may have yielded some interesting results. However, while missing data is a common problem in meta-analytic research, it should not preclude one from conducting a quantitative review. In fact, one of the reasons for conducting a meta-analysis is to identify areas of weakness and provide directions for future research. With the former in mind, we believe that future studies should include, and editors publish, complete information regarding whether any drugs were taken that could enhance BMD, cigarette smoking, diet, previous physical activity habits, alcohol consumption, and previous fractures. In addition, future studies should probably assess and report the different ground-reaction forces associated with the physical activity interventions they employ. We believe that this is critical to the establishment of more precise guidelines aimed at enhancing BMD.

In conclusion, the overall results of this study suggest that aerobic exercise has a small but positive effect on BMD at the lumbar spine and femur in women. ■

This study was funded by the United States Department of Defense, Army Medical Research and Materiel Command Award #17-98-1-8513.

REFERENCES

1 National Osteoporosis Foundation. *1996 and 2015 Osteoporosis Prevalence Figures: State by State Report.* Washington, DC: National Osteoporosis Foundation; 1997.

2 US Bureau of the Census Current Population Reports. *Population Projections of the United States by Age, Sex, and Hispanic Origin: 1995–2050.* Washington, DC: US Bureau of the Census; 1996:25–1130.

3 Ray NF, Chan JK, Thamer M, et al. Medical expenditures for the treatment of osteoporotic fractures in the United States in 1995: report from the National Osteoporosis Foundation. *J Bone Miner Res.* 1997;12:24–35.

4 Ernst E. Exercise for female osteoporosis: a systematic review of randomised clinical trials. *Sports Med.* 1998;25:359–368.

5 Kelley GA. Aerobic exercise and lumbar spine bone mineral density in postmenopausal women: a meta-analysis. *J Am Geriatr Soc.* 1998;46:143–152.

6 Kelley GA. Aerobic exercise and bone density at the hip in postmenopausal women: a meta-analysis. *Prev Med.* 1998;27:798–807.

7 Kelley GA. Exercise and regional bone mineral density in postmenopausal women: a meta-analytic review of randomized trials. *Am J Phys Med Rehabil.* 1998;77:76–87.

8 Egger M, Davey Smith G. Meta-analysis: bias in location and selection of studies. *BMJ.* 1998;316:61–66.

9 Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis.* San Diego, CA: Academic Press; 1985.

10 Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* New York, NY: Academic Press; 1988.

11 Follmann D, Elliot P, Suh I, et al. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol.* 1992;45:769–773.

12 Light RJ, Pillemer DB. *Summing up: The Science of Reviewing Research.* Cambridge, MA: Harvard University Press; 1984.

13 Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17:1–12.

14 Bloomfield SA, Williams N, Lamb D, et al. Nonweightbearing exercise may increase lumbar spine bone mineral density in healthy postmenopausal women. *Am J Phys Med Rehabil.* 1993;72:204–209.

15 Blumenthal JA, Emery CF, Madden DJ, et al. Effects of exercise training on bone density in older men and women. *J Am Geriatr Soc.* 1991;39:1065–1070.

16 Bouxsein ML. *Physical Activity and Bone Density.* Palo Alto, CA: Stanford University; 1992.

17 Brooke-Wavell KSF, Jones PRM, Hardman AE. Brisk walking reduces calcaneal bone loss in post-menopausal women. *Clin Sci (Colch).* 1997;92:75–80.

18 Caplan GA, Ward JA. The benefits of exercise in postmenopausal women. *Aust J Public Health.* 1993; 17:23–26.

19 Cavanaugh DJ, Cann CE. Brisk walking does not stop bone loss in postmenopausal women. *Bone.* 1988;9:201–204.

20 Dalsky GP, Stocke KS, Ehsani AA, et al. Weight-bearing exercise training and lumbar bone mineral content in postmenopausal women. *Ann Intern Med.* 1988;108:824–828.

21 Ebrahim S, Thompson PW, Baskaran V, et al. Randomized placebo-controlled trial of brisk walking in the prevention of postmenopausal osteoporosis. *Age Ageing.* 1997;26:253–260.

22 Grove KA, Londeree BR. Bone density in postmenopausal women: high impact versus low impact exercise. *Med Sci Sports Exerc.* 1992;24:1190–1194.

23 Hatori M, Hasegawa A, Adachi H, et al. The effects of walking at the anaerobic threshold level on vertebral bone loss in postmenopausal women. *Calcif Tissue Int.* 1993;52:411–414.

24 Heinonen A, Kanus P, Sievanen H, et al. Randomised control trial of effect of high-impact exercise on selected risk factors for osteoporotic fractures. *Lancet.* 1996;348:1343–1347.

25 Heinonen A, Oja PSH, Pasanen M, et al. Effect of two training regimens on bone mineral density in healthy perimenopausal women: a randomized controlled trial. *J Bone Miner Res.* 1998;13:483–490.

26 Kohrt WM, Snead DB, Slatopolsky E, et al. Additive effects of weight-bearing exercise and estrogen on

continued on page 452

12 Bjarnason I, Smethurst P, Clark P, et al. Effect of prostoglandin on indomethacin-induced increased intestinal permeability in man. *Scand J Gastroenterol.* 1989;164(suppl):97–102.

13 Tanaka S. An epidemiological survey on food-dependent exercise-induced anaphylaxis in kindergartners, schoolchildren, and junior high school students. *Asia Pac J Public Health.* 1994;7(1):26–30.

14 Schwartz LB, Metcalfe DD, Miller JS, et al. Tryptase levels as an indicator of mast-cell activation in systemic anaphylaxis and mastocytosis. *N Engl J Med.* 1987;316:1622–1626.

15 Simons FER, McMillan JL, Simons KJ. A double-blind, single-dose, crossover comparison of cetirizine, terfenadine, loratidine, astemizole, and chlorpheniramine versus placebo: suppressive effects on histamine-induced wheals and flares during 24 hours in normal subjects. *J Allergy Clin Immunol.* 1990;86:540–547.

16 Fujimoto S, Kurihara N, Hirata K, et al. Successful prophylaxis of wheat-dependent exercise-induced anaphylaxis with terfenadine. *Intern Med.* 1995;34(7): 654–656.

17 Harvey RP, Wegs J, Schocket AL, et al. A controlled trial of therapy in chronic urticaria. *J Allergy Clin Immunol.* 1981;68(99):675–679.

18 Margolis CF, Nisi R. Urticaria in a family practice. *J Fam Pract.* 1985;20:57–64.

19 Katsunuma T, Iikura Y, Akasawa A, et al. Wheat-dependent exercise-induced anaphylaxis inhibition by sodium bicarbonate. *Ann Allergy.* 1992;68:184–188.

20 Juji F, Suko M. Effectiveness of disodium cromoglycate in food-dependent, exercise-induced anaphylaxis: a case report. *Ann Allergy.* 1994;72:454.

21 Orfan NA, Stoloff RS, Harris KE, et al. Idiopathic anaphylaxis: total experience with 225 patients. *Allergy Proc.* 1992;13:35–43.

Kelley (continued from page 433)

bone mineral density in older women. *J Bone Miner Res.* 1995;10:1303–1311.

27 Kohrt WM, Ehsani AA, Birge SJ. HRT preserves increases in bone mineral density and reductions in body fat after a supervised exercise program. *J Appl Physiol.* 1998;84:1506–1512.

28 Kohrt WM, Ehsani AA, Birge SJ. Effects of exercise involving predominantly either joint-reaction or ground-reaction forces on bone mineral density in older women. *J Bone Miner Res.* 1997;12:1253–1261.

29 Lau EMC, Woo J, Leung PC, et al. The effects of calcium supplementation and exercise on bone density in elderly Chinese women. *Osteoporos Int.* 1992;2:168–173.

30 Little KD. *Effect of Exercise Mode on Bone Mineral Mass in Recently Postmenopausal Women.* Kent, OH: Kent State University; 1992.

31 Martin D, Notelovitz M. Effects of aerobic training on bone mineral density of postmenopausal women. *J Bone Miner Res.* 1993;8:931–936.

32 McMurdo MET, Mole PA, Paterson CR. Controlled trial of weight-bearing exercise in older women in relation to bone density and falls. *BMJ.* 1997;314:569.

33 Nelson ME, Fisher EC, Dilmanian FA, et al. A 1-y walking program and increased dietary calcium in postmenopausal women: effects on bone. *Am J Clin Nutr.* 1991;53:1304–1311.

34 Peterson SE, Peterson MD, Raymond G, et al. Muscular strength and bone density with weight training in middle-aged women. *Med Sci Sports Exerc.* 1991;23:499–504.

35 Prince R, Devine A, Criddle A, et al. The effects of calcium supplementation (milk powder or tablets) and exercise on bone density in postmenopausal women. *J Bone Miner Res.* 1995;10:1068–1075.

36 Prince RL, Smith M, Dick IM, et al. Prevention of postmenopausal osteoporosis: a comparative study of exercise, calcium supplementation, and hormone replacement therapy. *N Engl J Med.* 1991;325: 1189–1195.

37 Rikli RE, McManis BG. Effects of exercise on bone mineral content in postmenopausal women. *Res Q Exerc Sport.* 1990;61:243–249.

38 Ryan AS, Nicklas BJ, Dennis KE. Aerobic exercise maintains regional bone mineral density during weight loss in postmenopausal women. *J Appl Physiol.* 1998;84:1305–1310.

39 Snow-Harter C, Bouxsein ML, Lewis BT, et al. Effects of resistance and endurance exercise on bone mineral status of young women: a randomized exercise intervention trial. *J Bone Miner Res.* 1992;7: 761–769.

40 Tsukahara N, Toda A, Goto J et al. Cross-sectional and longitudinal studies on the effect of water exercise in controlling bone loss in postmenopausal women. *J Nutr Sci Vitaminol (Tokyo)* 1994;40:37–47.

41 Karpf DB, Shapiro DR, Seeman E et al. Prevention of nonvertebral fractures by alendronate: a meta-analysis. *JAMA.* 1997;277:1159–1164.

42 Kanders B, Dempster DW, Lindsay R. Interaction of calcium nutrition and physical activity on bone mass in young women. *J Bone Miner Res.* 1988;3: 145–149.

43 Bassey EJ. Mechanical loading and skeletal integrity: much needed debate. *J Bone Miner Res.* 1999;14: 1014–1015.

44 Sterling T D, Rosenbaum WL, Weinkam JJ. Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa. *Am Statistician.* 1995;49:108–112.

45 Cook DJ, Guyatt GH, Ryan G et al. Should unpublished data be included in meta-analyses? Current convictions and controversies. *JAMA.* 1993;269(21): 2749–2753.

46 Petitti DB. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine.* 2nd ed. New York, NY: Oxford University Press; 2000.

APPENDIX B

Retrieval of Individual Patient Data for an Exercise Meta-Analysis

George A. Kelley, DA;[1] Kristi S. Kelley, MEd;[1] Zung Vu Tran, PhD[2]
Clinical Investigation Program, Massachusetts General Hospital Institute of Health Professions, Boston, MA;[1] and the Department of Preventive Medicine and Biometrics, University of Colorado Health Sciences Center, Denver, CO[2]

The purpose of this study was to examine the feasibility of acquiring individual patient data (IPD) for a meta-analysis on the effects of exercise on bone mineral density in adults. We were able to obtain data from 29 (38.2%) of the 76 eligible studies. Binary multiple logistic regression analysis revealed a trend suggesting that authors of studies conducted in the United States were less likely to supply IPD when compared with authors who conducted studies in other countries (adjusted odds ratio, 0.324; 95% confidence interval, 0.104–1.004). Only 19% of authors from studies conducted in the United States vs. 52.9% of authors from other countries provided us with IPD. We conclude that we received a low response rate in the acquisition of IPD for a meta-analysis dealing with the effects of exercise on bone mineral density in adults. The use of summary means vs. IPD may be more appropriate for studies of this nature. (Am J Med Sports. 2002;4:350–354)

Address for correspondence:
George A. Kelley, DA, Associate Professor, Graduate Program in Clinical Investigation, Director, Meta-Analytic Research Group, MGH Institute of Health Professions, Charlestown Navy Yard, 36 First Avenue, Boston, MA 02129
E-mail: gakelley@bics.bwh.harvard.edu
Manuscript received September 18, 2001; accepted December 14, 2001

The use of meta-analysis is becoming increasingly common in the exercise training and physical activity literature. A recent MEDLINE search by one of the authors (KSK) found that the number of citations listed using the keywords "exercise and meta-analysis" has increased from two between the years 1980–1985 to 121 between the years 1995–2000 (unpublished results). To the best of our knowledge, all meta-analyses on this topic conducted to date have derived their results from the aggregation of summary data provided in the studies. An alternative approach is the retrieval of individual patient data (IPD) from study authors.

One of the major advantages of using IPD in a meta-analysis is the potential for increased statistic power as well as a more thorough examination of potential covariates.[1–3] Therefore, the use of IPD may be especially appropriate, since many meta-analyses include a small number of studies, thus limiting the interpretation and application of the findings.

One of the potential disadvantages with the retrieval of IPD is the inability to obtain IPD from studies that meet one's predefined inclusion criteria. This results in a form of bias known as retrieval bias.[2] In addition, meta-analyses of IPD are traditionally more expensive and labor-intensive than meta-analyses using summary means. Consequently, the use of summary data from individual studies may be preferable.

We have previously published meta-analytic work dealing with the effects of exercise on bone mineral density (BMD) in adult humans.[4–8] While our previous work has resulted in some noteworthy findings, these meta-analyses were based on the aggregation of summary data from individual studies. Unfortunately, we were limited in our ability to perform subgroup analyses because of a lack of information. The ability to examine potential factors associated with exercise-induced changes in BMD is important for deriving a better understanding of the true relationship between exercise and BMD. Since the acquisition of IPD could lead to a more accurate determination of the role of exercise on BMD, we sought such data from study

authors. Thus, the purpose of this paper is to report on the level of success of acquiring IPD dealing with the effects of exercise on BMD in adult humans.

ACQUISITION OF IPD DATA. The acquisition of IPD was conducted according to the general guidelines of Friedenreich.[9] For this study, references for IPD were derived from a database that contained 76 studies that met our previously defined meta-analytic inclusion criteria on the effects of exercise on BMD in adults (references available on request). Prior to sending out our request for IPD, a cover letter and IPD request sheet were developed, reviewed, revised, and approved by the three authors. We then sent, via postal mail, a copy of the cover letter and an IPD data acquisition form to the corresponding authors of the 76 studies. A follow-up request, approximately 5 weeks later, was sent to all authors who did not respond to our initial request. If the corresponding author referred us to one of the coauthors, contact was made with that author in an attempt to retrieve IPD. The first request contained no deadline date for the receipt of IPD. However, the second request included a deadline date of approximately 4 weeks from the date of mailing for the receipt of IPD. This deadline was extended for those authors who contacted us to request additional time to provide us with IPD. Some individual patient data were already available from five of the original studies in our database (i.e., from the published tables). However, requests were also sent to the corresponding authors of these studies in the event that additional IPD data might be provided. All authors who supplied IPD were mailed a check for US $40 to help cover incurred costs. We were limited to this amount of money because of budget limitations. Prior to the start of this study, approval was obtained by the Institutional Review Board at Massachusetts General Hospital.

STATISTICAL ANALYSIS. Descriptive statistics (frequencies, percentages, ranges, means, and standard deviations) were used to report overall results. Binary multiple logistic regression was used to examine potential predictors for whether IPD were finally sent to us or not. Predictors in the model included gender of author contacted, source of publication (journal vs. other), country in which the study was conducted (United States vs. other), and year of publication. The likelihood ratio statistic and Hosmer and Lemeshow test were used to identify whether the model adequately fit the data. The Nagelkerke *R*-squared statistic was used to identify the amount of variance accounted for by the predictor variables.[10] The Nagelkerke *R*-squared statistic is an adjusted version of the Cox and Snell *R*-squared. This adjustment was necessary because the Cox and Snell *R*-squared statistic has a value less than 1 even for a perfect model. Significance of regression coefficients for individual predictor variables was examined using the Wald statistic. In addition, odds ratios and 95% confidence intervals (CI), adjusted for other variables in the model, were used to examine the significance of individual predictor variables. Comparison of models with and without interactions was examined using the G test, which compares the log likelihoods between two models.

For comparative purposes, BMD results were calculated using the standardized difference effect size (ES) estimated from the summary data reported in the studies and corrected for small sample bias.[11] We were unable to use the original metric because of missing data. The standardized difference ES was calculated by taking the difference in BMD between the exercise and control groups and dividing by the pooled standard deviation of the exercise and control groups.[11] For studies that did not supply these data, the ES was calculated from other reported statistics by previously developed methods.[12] In general, an ES of 0.20 is considered a small effect, 0.50 a moderate effect, and 0.80 a large effect.[13] An ES of 0.20 for example, means that the exercise group differed from the control group by only two tenths of a standard deviation in favor of the exercise group. We then compared ES differences between those studies that did and did not provide IPD, using an analysis of variance-like random effects model developed for meta-analytic research.[11] This was accomplished by examining the between (Q_b) and within (Q_w) group differences for the ESs and their variances from each group. In addition, for those studies that supplied IPD, this approach was used to examine whether any statistically significant or clinically important differences existed between calculations of ESs from IPD and summary data reported in the studies.

The alpha level for a type I error was set at $p \leq 0.05$. Ninety-five percent CIs that did not cross zero were considered statistically significant. Trends were defined as values >0.05 but ≤ 0.10.[14,15]

Results

DESCRIPTION OF RESPONSES. Of the 76 requests mailed out, 41 (53.9%) authors responded, 33 (43.4%) did not respond at all, and two (2.6%) were returned to us because of undeliverable/invalid addresses. The reasons given by those authors who responded to our request but never supplied IPD data are shown in Table 1. Of the 41 who did respond, 26

of the 74 total, or 35.1%, provided us with IPD data. Of the 26 authors who provided IPD data, 22 (84.6%) sent their data as an attachment via electronic mail (our suggested preference), while two each sent data via either postal mail (7.7%) or facsimile (7.7%). The time taken from the date initial letters of request were mailed to the date that data were received ranged from 14–89 days ($\bar{X}$ ±SD = 50±23 days). Of the 22 authors who provided IPD via electronic mail, 15 (68.2%) included their data as a Microsoft Excel attachment (our suggested preference), while the remaining seven (31.8%) provided data as an SPSS file. Individual patient data provided from one author (for one study) could not be used because of missing data for BMD and our inability to contact this author at follow-up. Individual patient data from another author (for one study) was also excluded because it was a subset of data from another study already included in our database. Thus, we received usable IPD data for 24 of 76 studies (31.6%) for which data were requested. In addition, we already had in our possession IPD data from a total of five (6.6%) other studies. This left us with 29 studies (38.2%) for future IPD level analysis.

LOGISTIC REGRESSION ANALYSIS. The results of our binary multiple logistic regression analysis are shown in Table II. Approximately 21% of the variance was accounted for by the predictor variables (R^2_{adj}=0.207). Both the likelihood ratio statistic (χ^2=12.046; p=0.017) and Hosmer and Lemeshow test (χ^2=4.660; p=0.793) demonstrated that the model adequately fit the data. There was a trend for country where the study was conducted (United States vs. other) to be a predictor of whether or not IPD were provided, suggesting that authors of studies conducted in the United States were less likely to supply IPD when compared with authors who conducted studies in other countries. Only 19.0% of authors from the United States, vs. 52.9% of authors from other countries, provided us with IPD. No other variables were significant predictors for whether IPD would be provided.

Since there was a statistically significant association between country and year of publication (r=0.330; p=0.004), we compared our original model with a second model that included the interaction between country and year of publication. No statistically significant difference was found between the two models (G=0.464; p=0.496).

ES COMPARISONS. No statistically significant or clinically important differences were found in BMD between those studies that provided IPD vs. those that did not (IPD provided: $\bar{X}$±SD=0.134±0.364; 95% CI=0.069–0.198; IPD not provided: $\bar{X}$±SD=0.195±0.387; 95% CI=0.143–0.247; Q_b=2.126; p=0.145). In addition, for those studies that supplied IPD, no statistically significant or clinically important differences were found between calculations of ES from IPD and summary data reported in the studies (calculations from IPD: $\bar{X}$±SD=0.179±0.413; 95% CI=0.106–0.252; calculations summary data: $\bar{X}$±SD=0.134±0.364, 95% CI=0.069–0.198; Q_b=0.664; p=0.415).

Discussion

While the acquisition of IPD for meta-analytic purposes can lead to increased statistical power and a more thorough examination of potential covariates, the results of our investigation suggest that obtaining such data from authors of intervention studies dealing with the effects of exercise training on BMD in adults is difficult (31.6% of authors contacted provided IPD). This, coupled with the increased costs associated with the retrieval of IPD,[1] as well as the fact that we found no differences in BMD between studies that provided IPD vs. those that did not, suggests that the use of summary data from the actual studies may be more appropriate for examining the effects

Table I. Responses of Investigators Who Responded to the Authors' Request but Never Supplied IPD

No. of Authors	Results/Responses
6	Data no longer available
2	Data no longer available because of a change in computer systems
3	Corresponding author did not have data; referred us to another author who did not respond and/or did not send data
1	Expressed an interest in providing data but never provided such
1	Did not have time to track down data
1	Not willing to supply data until published in a refereed journal (original source was a dissertation)
1	Not willing to supply data because meta-analysis is inappropriate for exercise and bone mineral density studies

Table II. Results of Multiple Logistic Regression Analysis (n=74)

VARIABLE	B	SE	*DF*	SIGNIFICANCE	EXP(B)	95% CI
Constant	-275.443	190.997	1	0.149	0.000	NA
Country	-1.128	0.577	1	0.051*	0.324	0.104–1.004*
Gender	0.246	0.544	1	0.652	1.279	0.440–3.716
Source	0.614	1.179	1	0.602	1.848	0.183–18.627
Year	0.138	0.096	1	0.146	1.150	0.951–1.385

B=regression coefficients for the logistic regression; SE=standard error of the regression coefficients; *df*=degrees of freedom; Exp(B)=odds ratio, adjusted for independent variables; 95% CI=95% confidence interval for the odds ratio; *trend for statistical significance

of exercise training on BMD in adults. In addition, the inability to acquire IPD results in greater information bias, thus limiting the interpretation of findings from such studies. Thus, the use of summary means vs. IPD when conducting a meta-analysis on the effects of exercise training on BMD in humans should result in a more accurate as well as cost- and time-effective investigation. This is important because it would allow those with limited resources to conduct studies of this nature. Furthermore, given the proliferation of information in the health care field, a continued need for meta-analysis will exist.

We are not aware of anyone else who has attempted to retrieve IPD for an exercise-related meta-analysis. While the retrieval of IPD for meta-analyses may be problematic across all fields, including exercise, our results suggest that it may be especially problematic for those individuals interested in conducting IPD meta-analyses of exercise and bone studies. For example, Arnot et al.[16] were able to retrieve IPD from five of seven trials (71.4%) dealing with the effects of preoperative radiation therapy in esophageal carcinoma. Another meta-analysis reported the retrieval of IPD from 39 of 63 studies (61.9%) that met their inclusion criteria on the topic of breast cancer and hormone replacement therapy.[17] This compares to approximately 32% in our study.

One of the surprising findings of this study was the trend for more authors from studies conducted in countries other than the United States to provide us with IPD. While purely speculative, it may be that authors of studies conducted in the United States were less likely to provide us with IPD because they did not want to take the time to retrieve such information. This may be especially true given the small amount of money (US $40) we provided them for the retrieval of such. Unfortunately, we were not able to provide more money because of budget limitations. Alternatively, $40 to researchers in some foreign countries may represent a significant amount of money, thus resulting in a greater willingness to supply IPD.

Another possible reason for the low response rate from studies conducted in the United States may have to do with the investigators' concerns about protecting their data because of the potential misuse of such. For example, the strict guidelines that are enforced by the vast majority of university and hospital institutional review boards in the United States surrounding issues such as subject confidentiality may have precluded authors from supplying us with IPD. However, we believe that concerns about approbation from institutional review boards should not be an issue, as researchers should have data storage systems that protect the confidentiality of patients. Consequently, the sharing of IPD should not be a problem. Researchers who informed us that the data are "no longer available" were troubling in that the failure to sustain IPD in a manner that allows for verification of an analysis might be considered an ethical issue. Alternatively, this may be an issue of nothing more than selfishness on the part of some investigators. Since cooperation and trust are part of the foundation of science, we believe that any acts of selfishness on the part of investigators should be discouraged.

Since our investigation was limited to studies dealing with the effects of exercise on BMD, it may be inappropriate to generalize our results to other exercise meta-analyses. This, coupled with the fact that we are not aware of any other work in the meta-analytic field that has focused on predictors for retrieval of IPD, would lead us to suggest that future research in the meta-analytic field in general, and the exercise and meta-analysis field in particular, focus on this area. This may be especially important, since only 21% of the variance was accounted for in our logistic regression model. Thus, it appears that there may be other unknown factors, or combinations of factors, surrounding the retrieval of IPD.

In conclusion, the results of our study suggest a low response rate in acquiring IPD for a meta-analysis dealing with the effects of exercise training on BMD in adults, and show that success appears to be greater when IPD are requested for studies conducted in countries other than the United States. Given the relatively low response rate, and thus increased bias, the use of summary data may be more appropriate for examining the effects of exercise training on BMD in adults. ■

Acknowledgments: This study was supported by a grant from the Department of Defense, United States Army, Medical Research and Material Command Award #17-98-1-8513 (G.A. Kelley, Principal Investigator).

The authors would like to thank the following people for taking the time and effort to contribute their individual patient data to this project: ***K. Bennell, S. Bloomfield, G. Bravo, K. Brooke-Wavell, G. Caplan, I. Dick, S. Ebrahim, P. Gleeson, A. Heinonen, J. Iwamota, A.C. Kronhed, K. Little, S. Lord, D. Martin, N. Nara-Ashizawa, D. Nichols, E. Preisinger, R. Prince, J. Pritchard, A. Ryan, G. Smidt, and K. Uusi-Rasis.***

REFERENCES

1 Clarke M, Godwin J. Systematic reviews using individual patient data: a map for the minefields? *Ann Oncol.* 1998;9:827–833.
2 Lau J, Ioannidis JPA, Schmid CH. Summing up evidence: one answer is not always enough. *Lancet.* 1998;351:123–127.
3 Petitti DB. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine.* 2nd ed. New York, NY: Oxford University Press; 2000:136–139.
4 Kelley GA. Aerobic exercise and bone density at the hip in postmenopausal women: a meta-analysis. *Prev Med.* 1998;27:798–807.
5 Kelley GA. Aerobic exercise and lumbar spine bone mineral density in postmenopausal women: a meta-analysis. *J Am Geriatr Soc.* 1998;46:143–152.
6 Kelley GA. Exercise and regional bone mineral density in postmenopausal women: a meta-analytic review of randomized trials. *Am J Phys Med Rehabil.* 1998;77:76–87.
7 Kelley GA, Kelley KS, Tran ZV. Exercise and bone mineral density in men: a meta-analysis. *J Appl Physiol.* 2000;88:1730–1736.
8 Kelley GA, Kelley KS, Tran ZV. Resistance training and bone mineral density in women: a meta-analysis of controlled trials. *Am J Phys Med Rehabil.* 2001;80:65–77.
9 Friedenreich CM. Methods for pooled analyses of epidemiologic studies. *Epidemiology.* 1993;4:295–302.
10 Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika.* 1991;78(3):691–692.
11 Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis.* San Diego, CA: Academic Press; 1985:13, 78, 122–130, 149–165, 167–183.
12 Johnson BT. DSTAT: Software for the meta-analytic reviews of research literatures. Version 1.10. Hillsdale, NJ: Lawrence Erlbaum Associates; 1993:51–62.
13 Cohen J. A power primer. *Psychol Bull.* 1992;112:155–159.
14 Norman GR, Streiner DL. *Biostatistics: The Bare Essentials.* Hamilton, Ontario: B.C. Decker; 2000:46.
15 Rosnow RL, Rosenthal R. Statistical procedures and the justification of knowledge in psychological science. *Am Psychol.* 1991;44:1276–1284.
16 Arnot SJ, Duncan W, Gignoux M, et al. Preoperative radiotherapy in esophageal carcinoma: a meta-analysis using individual patient data. *Int J Radiat Oncol.* 1998;41:579–583.
17 Collaborative group on hormonal factors in breast cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiologic studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet.* 1997;350:1047–1059.

APPENDIX C

Journal of Gerontology: MEDICAL SCIENCES
2002, Vol. 57A, No. 9, M599–M604

Exercise and Lumbar Spine Bone Mineral Density in Postmenopausal Women: A Meta-Analysis of Individual Patient Data

George A. Kelley,[1] Kristi S. Kelley,[1] and Zung Vu Tran[2]

[1]Clinical Investigation Program, MGH Institute of Health Professions, Boston, Massachusetts.
[2]Department of Preventive Medicine and Biometrics, University of Colorado Health Sciences Center, Denver.

Background. Low bone mineral density (BMD) at the lumbar spine is a major public health problem among postmenopausal women. We conducted a meta-analysis of individual patient data (IPD) to examine the effects of exercise on lumbar spine BMD in postmenopausal women.

Methods. IPD were requested from a previously developed database of summary means from randomized and nonrandomized trials dealing with the effects of exercise on BMD. Two-way analysis of variance tests with pairwise comparisons ($p \leq .05$) and 95% confidence intervals (CIs) were used to determine the statistical significance for changes in lumbar spine BMD.

Results. Across 13 trials that included 699 subjects (355 exercise, 344 control), a statistically significant interaction was found between test and group ($F = 15.232$, $p = .000$). Pairwise comparisons (Bonferroni t tests) revealed a statistically significant increase in final minus initial BMD for the exercise group ($\bar{X} \pm SD = 0.005 \pm 0.043$ g/cm^2, $t = 2.46$, $p = .014$, 95% CI = 0.001–0.009) and a statistically significant decrease in final minus initial BMD for the control group ($\bar{X} \pm SD = -0.007 \pm 0.045$ g/cm^2, $t = -3.051$, $p = .002$, 95% CI = −0.012–−0.002). Changes were equivalent to an approximate 2% benefit in lumbar spine BMD (exercise, +1%, control, −1%).

Conclusions. The results of this IPD meta-analysis suggest that exercise helps to improve and maintain lumbar spine BMD in postmenopausal women.

IT has been estimated that approximately 26.2 million postmenopausal women have either osteoporosis or osteopenia (1). As a result of having osteoporosis or osteopenia, a person is at an increased risk for fracture, particularly at the vertebrae, hip, and distal forearm (2). Of these three sites, fractures of the vertebrae, which represent approximately 56% of all fractures, are the most common, with an estimated 700,000 per year (2). The health-care costs associated with vertebral fractures were estimated to be approximately $746 million in 1995 and are expected to increase substantially in the future (3).

One of the potential interventions for increasing and/or maintaining vertebral bone mineral density (BMD) in postmenopausal women is exercise, a low-cost, nonpharmacologic intervention that is available to most individuals. We have recently conducted meta-analytic work in which we reported improvements in lumbar spine BMD because of exercise in postmenopausal women (4,5). This work was based on the most commonly used approach for conducting meta-analytic work, that is, the abstraction of summary means from studies meeting specified inclusion criteria. However, the use of individual patient data (IPD) versus summary means from eligible studies represents the most comprehensive approach for conducting meta-analytic work, including the potential for increased statistical power as well as a more thorough examination of potential covariates (6–8). Given the health-care consequences of low BMD at the lumbar spine, the possible benefit of exercise for improving and/or maintaining lumbar spine BMD, and the potential for a meta-analysis of IPD to provide more thorough information regarding the effect of exercise on lumbar spine BMD, we sought to examine the effects of exercise on lumbar spine BMD in postmenopausal women by conducting a meta-analysis using IPD.

Methods

Data Sources

From a previously developed meta-analytic database that included the summary means from 76 studies dealing with the effects of exercise on BMD, we sought to obtain IPD. Briefly, IPD were requested by sending a cover letter and data request sheet to authors via postal mail. For those who did not respond to our initial request, a follow-up letter was sent via postal mail approximately 5 weeks later.

Study Selection

From the database of 76 studies, we included studies that met the following criteria: (i) randomized and nonrandomized trials that included a comparative control (nonexercise) group, (ii) exercise lasting at least 16 weeks, (iii) postmenopausal women only, (iv) journal articles, dissertations, and masters theses published in the English-language literature, (v) studies published between January 1966 and December

1998, (vi) BMD (relative value of bone mineral per measured bone area) assessed at the lumbar spine, and (vii) ability to obtain IPD from authors. Despite the fact that methods to assess BMD (dual-photon absorptiometry [DPA], dual-energy x-ray absorptiometry [DEXA]) have only been widely available since the 1980s, we searched back to 1966 to ensure that there was no comparative technology that we might have missed. We did not include studies from non-English-language journals because of the potential for error in the translation and interpretation of findings. If more than one study included the same subjects, for example, a dissertation and refereed journal article, we retrieved and referenced both studies to extract the maximum amount of information but only included this as one data set.

Data Abstraction

All data were abstracted on a coding sheet that could hold up to 91 pieces of information from each study. All data were coded and verified for accuracy and consistency by George A. Kelley. Blinding of the coder to the identity and institutional affiliation of the authors as well as study results was not performed because it has been shown that these procedures have neither a statistically significant nor a clinically important effect on the results (9). The major categories coded included study, subject, BMD assessment, and training program characteristics as well as primary and secondary outcomes.

Statistical Analysis

Initial subject characteristics.—Potential differences between initial subject characteristics for exercise and control groups were examined using independent t tests and 95% confidence intervals (CIs) for continuous variables and 2 × 2 chi-square tests for categorical variables.

Primary outcomes.—Initial and final values for lumbar spine BMD between exercise and control groups were examined by using a two-way analysis of variance (ANOVA) test with repeated measures on one factor (test). Because this was an unbalanced design, a General Linear Model was used. Pairwise comparison tests (Bonferroni t tests) were used to identify the specific location of the observed interaction between test (final vs initial) and group (exercise vs control). To examine for outliers, ANOVAs were performed with each study deleted from the model once. Because of missing data, we were unable to include potential covariates in the ANOVA model. Consequently, we used Pearson-Product moment correlations to examine for potential associations between changes in BMD and age, height, body weight, years postmenopausal, cigarette smoking, alcohol consumption, calcium and vitamin D intake, compliance (percentage of exercise sessions attended), length of training (weeks), type of BMD assessment (DEXA, DPA), and study design (randomized vs nonrandomized controlled trial). We were unable to partition the data according to the different types of exercise because of the various interventions employed.

Because of the inability to retrieve IPD from all eligible studies, we also examined whether our results differed between studies according to the availability of IPD. To include all eligible studies in the analysis, we used the standardized difference effect size (ES) calculated from the summary data reported in the studies and corrected for small sample bias (10). In general, an ES of 0.20 is considered a small effect, 0.50 a moderate effect, and 0.80 a large effect (11). An ES of 0.80, for example, means that the exercise group differed from the control group by eight-tenths of a standard deviation in favor of the exercise group. We then compared ES differences between those studies in which IPD were provided versus those in which they were not using an ANOVA-like random effects model developed for meta-analytic research (10). This was accomplished by examining the between (Q_b) and within (Q_w) group differences for the ESs and their variances from each group.

Secondary outcomes.—Secondary outcomes (body weight, calcium intake, and vitamin D intake) were analyzed using the same ANOVA procedures that were used to evaluate changes in lumbar spine BMD. We used independent t tests to analyze initial differences between exercise and control groups for these variables because more data were available for initial values versus final values and we wanted to capture as much data as possible in our analyses.

Descriptive Statistics and Alpha Level

Means and standard deviations ($\bar{X} \pm SD$) were used to describe continuous variables, whereas frequencies and percentages were used for categorical variables. The alpha level for statistical significance was set at $p \leq .05$. Ninety-five percent CIs that did not cross 0.00 were also considered statistically significant.

Results

Study Characteristics

Of the 32 studies that met our criteria for inclusion, we were able to retrieve IPD from 13 (41%) (12–26). Note that the number of references exceeds the number of studies because two were published in dissertation (17,24) and two in journal format (18,25). A description of the studies is shown in Table 1. The 13 studies represented a total of 30 groups (17 exercise, 13 control) and 699 subjects (355 exercise, 344 control). Seven of the trials were randomized controlled trials, and the remaining six were nonrandomized controlled trials. The length of the studies ranged from 24 to 104 weeks ($\bar{X} \pm SD = 56 \pm 8$ weeks). Thirteen of the exercise groups included some type of weight-bearing exercise, two appeared to perform nonweight-bearing exercise, and the remaining two participated in weight training. Compliance, defined as the percentage of exercise sessions attended, averaged 75 ± 17%. Seven of the studies assessed lumbar spine BMD using DEXA, whereas the remaining six used DPA.

Initial Subject Characteristics

Initial subject characteristics for continuous and categorical variables can be found in Tables 2 and 3, respectively. For continuous variables, the number of years that the subjects were postmenopausal was significantly greater in the control versus exercise groups, whereas calcium intake was

Table 1. Characteristics of Bone Mineral Density Studies (gm/cm^2) in Which IPD Were Provided for Postmenopausal Women at the Lumbar Spine

Study	Design/Subjects	Exercise Intervention(s)	BMD Assessment
Bloomfield and colleagues (12)	CT that included 18 postmenopausal women assigned to either an exercise (n = 7; age = 62.1 ± 2.1 years) or control (n = 11; age = 60.0 ± 9.4 years) group.	32 weeks of training performed 3× per week for 50 minutes per session (15-minute warm-up, 30 minutes of stationary cycling, 5-minute cool-down) at 60–80% of maximal heart rate.	DPA (Lunar DP3, Lunar Radiation, Madison, WI) at L1–L4.
Bravo and colleagues (13)	RCT that included 106 women assigned to either an exercise (n = 44; age = 59.8 ± 5.9 years) or control (n = 62; age = 60 ± 6.3 years) group.	52 weeks of training performed 3× per week for 60–65 minutes per session. Exercise sessions consisted of a 10-minute warm-up, 25 minutes of rapid walking replaced with aerobic dance 1× per week, and 15 minutes of bench stepping at 60–70% of MHRR. This was followed by 10–15 minutes of resistance exercise.	DEXA (Lunar DPX, Lunar Radiation) at L2–L4.
Brooke-Wavell and colleagues (14)	RCT of 76 postmenopausal women assigned to either an exercise (n = 37; age = 65.0 ± 2.8 years) or control (n = 39; age = 64.2 ± 3.1 years) group.	52 weeks of training that consisted of self-monitored walking 3.5 times per week for 14.8 minutes per day for the first 12 weeks, followed by 20.4 minutes per day of walking, 4.8 days per week, for the remainder of the study.	DEXA (Lunar DPX, Lunar Radiation) at L2–L4.
Caplan and colleagues (15)	CT of 30 postmenopausal women assigned to either an exercise (n = 19; age = 66.4 ± 5.0 years) or control (n = 11; age = 65.4 ± 4.9 years) group.	104 weeks of aerobic weight-bearing exercise performed 2× week for 60 minutes (warm-up, 20–25 minutes of low-impact aerobic exercise, 10 minutes of ball games for improved hand-eye coordination followed by work on floor mats for strength and flexibility, 10 minutes of relaxation). Subjects were also asked to exercise on their own 1× per week so that the pulse would be elevated for at least 20–30 minutes.	DPA (Lunar DPA, Lunar Radiation).
Ebrahim and colleagues (16)	RCT of 92 postmenopausal women assigned to either an exercise (n = 47; age = 66.4 ± 7.9 years) or control (n = 45; age = 68.1 ± 7.8 years) group.	104 weeks of walking 3× per week for 40 minutes per session.	DEXA (Lunar DPX, Lunar Radiation).
Grove (17), Grove and Londeree (18)	RCT that included 15 postmenopausal women assigned to either a low-impact exercise group (n = 5; age = 56.6 ± 43.3 years), high-impact exercise group (n = 5; age = 54.0 ± 1.9 years), or control group (n = 5; age = 56.0 ± 4.5 years).	52 weeks of training performed 3× week for approximately 60 minutes per session (15–20 minute warm-up, 20 minutes of either low- or high-impact exercise, 15-minute cool-down). Low-impact activities were considered those that produced forces less than 1.5× body weight, high impact ≥2.0× body weight.	DPA (Lunar DP3, Lunar Radiation) at L2–L4.
Iwamoto and colleagues (19)	CT that included 35 postmenopausal women assigned to either an exercise (n = 15; age = 64.8 ± 6.1 years) or control (n = 20; age = 64.8 ± 5.7 years) group.	52 weeks of outdoor walking (7 days per week) and gymnastic training (at least 5 days per week).	DEXA (Norland XR26, Norland Medical Systems, White Plains, NY) at L2–L4.
Little (20)	CT that included 21 postmenopausal women assigned to a resistance training (n = 6; age = 59.5 ± 2.3 years), walking (n = 6; age = 52.3 ± 4.5 years), swimming (n = 5; age = 51.8 ± 5.8 years), or control (n = 4; age = 60.8 ± 1.4 years) group.	Resistance exercise consisted of 32 weeks of training with 9 exercises performed 3 times per week for 1 set of 8–12 repetitions at 60%–80% of 1RM; Walking consisted of 32 weeks of training, 3× per week for 30–50 minutes per session (5–10-minute warm-up; walking for 20–30 minutes; 5–10-minute cool-down) at 70%–90% of maximal heart rate; Swimming consisted of 32 weeks of training, 3× per week for 30–50 minutes per session (5–10-minute warm-up; walking for 20–30 minutes; 5–10-minute cool-down) at 70%–90% of maximal heart rate.	DPA (Lunar, Lunar Radiation) at L2–L4.
Lord and colleagues (21)	RCT that included 138 subjects assigned to either an exercise (n = 67; age = 70.8 ± 5.0 years) or control (n = 69; age = 71.0 ± 4.9 years) group.	42 weeks of exercise performed 2× per week for approximately 60 minutes per session (5-minute warm-up, 35–40 minutes of aerobic exercises [activities for balance, hand-eye and foot-eye coordination], strengthening exercises, 15 minutes of stretching, and 5–10 minute cool-down).	DEXA (Lunar DPX, Lunar Radiation) at L2–L4.
Martin and Notelovitz (22)	RCT that included 55 postmenopausal women assigned to a 30-minute exercise group (n = 20; age = 60.3 ± 7.8 years), 45-minute exercise group (n = 16; age = 57.8 ± 7.1 years), or control (n = 19; age = 56.7 ± 6.9 years) group.	52 weeks of treadmill exercise performed 3× week for either 30 or 45 minutes per session at 70–85% of maximal heart rate. Each session included a 3–5-minute warm-up and cool-down at 60% of maximal heart rate.	DPA (Lunar DP3, Lunar Radiation) at L2–L4.

Continued on next page

Table 1. Characteristics of Bone Mineral Density Studies (gm/cm²) in Which IPD Were Provided for Postmenopausal Women at the Lumbar Spine (*Continued*)

Study	Design/Subjects	Exercise Intervention(s)	BMD Assessment
Prince and colleagues (23)	RCT that included assignment of 63 postmenopausal women to a calcium and exercise (n = 35; age = 62.4 ± 4.8 years), or calcium only (n = 28; age = 63.2 ± 4.8 years) group.	104 weeks of weight-bearing exercise performed 2× week for approximately 60 minutes per session. Subjects were also asked to walk another 2 hours per week at 60% of peak heart rate for their age.	DEXA (QDR-1000, Hologic, Waltham, MA) at L1–L4.
Pruitt (24), Pruitt and colleagues (25)	CT that included 24 postmenopausal women assigned to either an exercise (n = 15; age = 53.6 ± 4.1 years) or control (n = 9; age = 55.6 ± 2.9 years) group.	36 weeks of strength training consisting of 13 exercises performed 3× week at 50%–60% of 1RM for 1 set of 10–12 repetitions for the upper body and 10–15 repetitions for the lower body.	DPA (Lunar DP3, Lunar Radiation) at L2–L4.
Ryan and colleagues (26)	CT that included 28 postmenopausal women assigned to either a weight loss (n = 15; age = 63.4 ± 5.7 years) or exercise + weight loss (n = 13; age = 61.3 ± 4.8 years) group.	24 weeks of aerobic exercise (treadmill jogging) performed 3× week for up to 35 minutes per session at 50 to >70% of $\dot{V}O_{2max}$. Each session included a 10-minute warm-up and cool-down period.	DEXA at L2–L4.

Notes: IPD = individual patient data; BMD = bone mineral density; CT = controlled trial; RCT = randomized controlled trial; DPA = dual-photon absorptiometry; DEXA = dual-energy x-ray absorptiometry; MHRR = maximal heart rate reserve; RM = repetition maximum. Study by Prince also included placebo and milk powder group but for comparison purposes, these groups were not included in our analysis. Only subjects who completed the study and for which BMD data were available are reported in the designs/subjects section; number of subjects reported as $\bar{X} \pm SD$. Bone density assessment limited to bone mineral density measures in g/cm².

greater in the exercise versus control groups. No statistically significant differences between exercise and control groups were observed for any other continuous or categorical variables.

Primary Outcomes

As can be seen in Table 4, there was an increase in lumbar spine BMD in the exercise groups and a decrease in the control groups. The mean difference between the two groups was 0.013 ± 0.079 g/cm², 95% CI = 0.007–0.019. These changes were equivalent to an approximate 2% benefit in lumbar spine BMD (exercise, +1%, control, −1%). The ANOVA results in Table 5 show a statistically significant main effect difference between group and an interaction between group and test. Pairwise comparison tests for the Group × Test interaction revealed a statistically significant increase in final versus initial BMD for the exercise groups (t = 2.464, p = .014), a statistically significant decrease in final versus initial BMD for control groups (t = −3.051, p = .002), and greater initial as well as final values for exercise groups compared to control groups (initial, t = 2.544, p = .011; final, t = 3.320, p = .001). Results were similar when each study was deleted from the model once. For the exercise groups, larger increases in lumbar spine BMD were associated with assessment of BMD using DEXA versus DPA (r = −0.126, p = .018, 95% CI = −0.227–−0.022). For control subjects, larger decreases in lumbar spine BMD were associated with younger age (r = 0.170, p = .002, 95% CI = 0.064–0.272), taller stature (r = −0.109, p = .048, 95% CI = −0.215–−0.005), absence of hormone replacement therapy (r = 0.152, p = .005, 95% CI = 0.047–0.254), assessment of BMD using DPA versus DEXA (r = −0.287, p = .000, 95% CI = −0.381–−0.187) and nonrandomized versus randomized controlled trials (r = 0.172, p = .001, 95% CI = 0.067–0.273). No other statistically significant or clinically relevant relationships were observed for the exercise or control groups. Finally, no statistically significant differences in lumbar spine BMD were found when we compared the 13 studies that included IPD (ES = 0.366 ± 0.423, 95% CI = 0.131–0.600) with the 19 studies that did not include IPD (ES = 0.219 ± 0.430, 95% CI = 0.059–0.379; Q_b = 1.184, p = .277).

Secondary Outcomes

No statistically significant main effects or interactions were found for body weight, calcium intake, or vitamin D intake.

Discussion

The primary purpose of meta-analysis is to reach some general conclusions about a body of research. The overall results of this study suggest that exercise helps to in-

Table 2. Initial Characteristics of Subjects for Continuous Variables

Variable	Exercise n	Exercise $\bar{X} \pm SD$	Control n	Control $\bar{X} \pm SD$	Significance t (p)	CI (95%)
Age (y)	340	63.9 ± 7.4	335	64.5 ± 7.4	−0.92 (.357)	−1.64 to 0.59
Height (cm)	329	158.9 ± 6.9	327	158.2 ± 7.0	1.18 (.239)	−0.43 to 1.71
Body Weight (kg)	340	65.1 ± 12.3	330	64.1 ± 13.0	1.04 (0.299)	−0.91 to 2.94
Postmenopause (y)	156	13.0 ± 9.9	147	17.3 ± 11.8	−3.46 (.001)*	−6.78 to −1.86*
Calcium (mg)	193	926.9 ± 394.0	195	834.7 ± 350.6	2.44 (.015)*	17.79 to 166.64*
Vitamin D (IUs)	55	195.9 ± 215.4	62	161.5 ± 132.4	1.05 (.294)	−30.27 to 99.13

Note: CI = confidence interval; IUs = international units.
*Statistically significant.

Table 3. Initial Characteristics of Subjects for Categorical Variables

Variable	Exercise n (%)	Control n (%)	χ_2 (p)
Cigarette Smoking	25 (9.9)	33 (12.7)	1.04 (.307)
Alcohol Consumption	130 (52.2)	121 (47.5)	1.14 (.285)
Estrogen/Progesterone Use	24 (6.9)	18 (5.4)	0.65 (.419)
Previous Fractures (any site)	61 (37.4)	62 (41.6)	0.57 (.450)
Race (white)	259 (94.1)	238 (91.2)	1.78 (.182)

Note: Results limited to studies that reported data for each variable.

Table 5. ANOVA Summary Table for Lumbar Spine BMD (General Linear Model)

Source of Variation	*df*	SS	*MS*	*F*	*p*	Partial η^2
Group	1	0.834	0.834	8.685	.003*	0.012
Error (Group)	697	66.962	0.096	—	—	—
Test	1	0.000	0.000	0.199	.656	0.000
Test × Group	1	0.001	0.001	15.232	.000*	0.021
Error (Test)	697	0.670	0.001	—	—	—

Note: SS = sum of squares; *MS* = mean square; Group = exercise vs control; Test = initial vs final.

*Significantly different, $p \leq .05$.

crease and maintain lumbar spine BMD in postmenopausal women. This supports our previous meta-analytic work of summary means and lumbar spine BMD (4,5), but is in contrast to our more recent meta-analytic work using IPD in which we found no statistically significant difference in femoral neck BMD (27). Although these are important findings, the clinical importance of an approximate 2% benefit, especially as it relates to fracture risk, cannot be elucidated at this time. However, although beyond the scope of this study, the increased strength, balance, and ambulatory skills that may be realized from a regular program of exercise may also help reduce the risk of falling and suffering subsequent fractures (28). Although we were unable to identify specific exercise programs for optimizing lumbar spine BMD, it would appear plausible to suggest that one adhere to the recent National Institutes of Health Consensus Statement that recommends participation in regular exercise, especially resistance and high-impact activities (28).

Our finding that larger decreases in BMD in the control groups were associated with younger age is not surprising given the fact that bone loss is most rapid during the early postmenopausal years (29). In addition, the observed association between the absence of hormone replacement therapy and greater decreases in lumbar spine BMD was also not surprising because hormone replacement therapy is an established therapeutic intervention for preserving BMD among postmenopausal women (28). However, we can offer no biological explanation regarding the observed association between greater decreases in lumbar spine BMD and taller stature. This is especially because it is generally believed that shorter women are considered more osteoporotic than taller women. Given this currently held notion, caution is warranted in the interpretation of this finding. Indeed, it may be that our observed association was nothing more than the play of chance given the large number of statistical tests that were conducted in our study.

Meta-analysis, like any other type of review, is limited by the availability of data and the limitations of the included studies. Thus, in addition to making the best of the existing data and trying to reach some overall conclusions regarding a body of research, it is also the meta-analyst's responsibility to identify areas of weakness to provide directions for future research.

For example, because we were unable to categorize the different types of exercise interventions, we would suggest that future researchers provide a better description of their exercise programs, especially as it relates to the forces employed during the exercise intervention. Consequently, exercise programs that provide optimal benefits to lumbar spine BMD can be recommended.

We were surprised that data on calcium intake were available for only 56% of the subjects included in this analysis. Because calcium intake is important for maintaining and/or increasing BMD in humans, it would seem reasonable to suggest that data on calcium intake be assessed and reported. In addition, because vitamin D intake is also important for the absorption of calcium and data on vitamin D intake were available for only 17% of the subjects included in this analysis, the assessment and reporting of this information also appears warranted.

Although white, non-Hispanic women are disproportionately affected with osteoporosis and low bone mass, the effect on other races is also significant. For example, the National Osteoporosis Foundation has reported that approximately 10% of black women older than 50 years have osteoporosis, and 29% have low bone mass. Additionally, 16% of American-Indian and Hispanic women aged 50 and older have osteoporosis, and 36% have low bone mass (30). Because approximately 93% of the subjects in this study were white and the responses to exercise in relation to BMD may vary by race, it is recommended that future studies include women from other ethnic groups.

Because data on the number of years that the subjects were postmenopausal were available for only 43% of the subjects included in this analysis, future research needs to include this type of information because it may be a potential confounder in relation to exercise-induced changes in lumbar spine BMD in postmenopausal women.

The fact that the vast majority of studies included in our meta-analysis were published in refereed journal articles may have led to an overestimate of the benefits of exercise on BMD at the lumbar spine because there is a tendency for authors to submit, and editors to publish, studies that yield statistically significant and positive results, i.e., publication bias (10).

Table 4. Lumbar Spine BMD Results (g/cm²)

Group	n	Initial ($\bar{X} \pm SD$)	Final ($\bar{X} \pm SD$)	Difference ($\bar{X} \pm SD$)	CI (95%)
Exercise	355	0.991 ± 0.221	0.996 ± 0.224	0.005 ± 0.043	0.001 ± 0.009*
Control	344	0.948 ± 0.218	0.941 ± 0.218	−0.007 ± 0.045	−0.012 ± −0.002*

Note: CI = confidence interval.

*Statistically significant.

For both exercise and control subjects, greater decreases in lumbar spine BMD were associated with assessment of BMD using DPA versus DEXA. Because DEXA is generally considered to be a more valid assessment of BMD and is currently the most common method used to assess BMD at the lumbar spine, the results from studies using DEXA may be more valid. The finding that greater decreases in lumbar spine BMD were associated with nonrandomized versus randomized trials suggests that randomized trials may yield more valid results.

Although the above-described associations are interesting, they should be viewed with caution for the following reasons: (i) they may have been nothing more than the play of chance given the large number of statistical tests that were conducted, (ii) we were unable to examine for potential interrelationships between variables because of missing data, and (iii) the associations accounted for only a small proportion of the total variance.

In conclusion, the results of this IPD meta-analysis suggest that exercise improves and maintains lumbar spine BMD in postmenopausal women.

ACKNOWLEDGMENTS

This study was supported by a grant from the Department of Defense, United States Army, Medical Research and Materiel Command Award #17-98-1-8513 (G.A. Kelley, Principal Investigator).

We thank the following people for taking the time and effort to contribute their individual patient data to this project: S. Bloomfield, G. Bravo, K. Brooke-Wavell, G. Caplan, S. Ebrahim, J. Iwamoto, K. Little, S. Lord, D. Martin, R. Prince, and A. Ryan.

Address correspondence to Dr. George A. Kelley, FACSM, Associate Professor, Graduate Program in Clinical Investigation, Director, Meta-Analytic Research Group, MGH Institute of Health Professions, Charlestown Navy Yard, 36 First Avenue, Boston, MA 02129. E-mail: gakelley@bics.bwh.harvard.edu

REFERENCES

1. Kanis JA, Melton LJ, Christiansen C, Johnston CC, Khaltaev N. The diagnosis of osteoporosis. *J Bone Miner Res.* 1994;9:1137–1141.
2. Riggs BL, Melton LJ. The worldwide problem of osteoporosis: insights afforded by epidemiology. *Bone.* 1995;17:505S–511S.
3. Ray NF, Chan JK, Thamer M, Melton LJ. Medical expenditures for the treatment of osteoporotic fractures in the United States in 1995: report from the National Osteoporosis Foundation. *J Bone Miner Res.* 1997;12:24–35.
4. Kelley GA. Aerobic exercise and lumbar spine bone mineral density in postmenopausal women: a meta-analysis. *J Am Geriatr Soc.* 1998;46:143–152.
5. Kelley GA, Kelley KS, Tran ZV. Resistance training and bone mineral density in women: a meta-analysis of controlled trials. *Am J Phys Med Rehabil.* 2001;80:65–77.
6. Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet.* 1993;341:418–422.
7. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. *Stat Med.* 1995;14:2057–2079.
8. Sylvester R, Collette L, Duchateau L. The role of meta-analyses in assessing cancer treatments. *Eur J Cancer.* 2000;36:1351–1358.
9. Berlin JA. Does blinding of readers affect the results of meta-analyses? *Lancet.* 1997;350:185–186.
10. Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis.* San Diego, Calif: Academic Press; 1985.
11. Cohen J. A power primer. *Psychol Bull.* 1992;112:155–159.
12. Bloomfield SA, Williams N, Lamb D, Jackson R. Non-weightbearing exercise may increase lumbar spine bone mineral density in healthy postmenopausal women. *Am J Phys Med Rehabil.* 1993;72:204–209.
13. Bravo G, Gauthier P, Roy PM, et al. Impact of a 12-month exercise program on the physical and psychological health of osteopenic women. *J Am Geriatr Soc.* 1996;44:756–762.
14. Brooke-Wavell KSF, Jones PRM, Hardman AE. Brisk walking reduces calcaneal bone loss in post-menopausal women. *Clin Sci (Colch).* 1997;92:75–80.
15. Caplan GA, Ward JA, Lord SR. The benefits of exercise in postmenopausal women. *Aust J Public Health.* 1993;17:23–26.
16. Ebrahim S, Thompson PW, Baskaran V, Evans K. Randomized placebo-controlled trial of brisk walking in the prevention of postmenopausal osteoporosis. *Age Ageing.* 1997;26:253–260.
17. Grove CA. *The Effects of High Impact Exercise Versus Low Impact Exercise on Bone Density in Postmenopausal Women* [dissertation]. Columbia, Mo: University of Missouri; 1990.
18. Grove KA, Londeree BR. Bone density in postmenopausal women: high impact versus low impact exercise. *Med Sci Sports Exerc.* 1992;24:1190–1194.
19. Iwamoto J, Takeda T, Otani T, Yabe Y. Effect of increased physical activity on bone mineral density in postmenopausal osteoporotic women. *Keio J Med.* 1998;47:157–161.
20. Little KD. *Effect of Exercise Mode on Bone Mineral Mass in Recently Postmenopausal Women* [dissertation]. Kent, Ohio: Kent State University; 1992.
21. Lord SR, Ward JA, Williams P, Zivanovic E. The effects of community exercise program on fracture risk factors in older women. *Osteoporos Int.* 1996;6:361–367.
22. Martin D, Notelovitz M. Effects of aerobic training on bone mineral density of postmenopausal women. *J Bone Miner Res.* 1993;8:931–936.
23. Prince R, Devine A, Criddle A, et al. The effects of calcium supplementation (milk powder or tablets) and exercise on bone density in postmenopausal women. *J Bone Miner Res.* 1995;10:1068–1075.
24. Pruitt L. *The Effect of an Isotonic Weight-Training Program Upon Bone Density and Selected Indices of Bone Turnover in Early Postmenopausal Women* [dissertation]. Columbus, Ohio: Ohio State University; 1989.
25. Pruitt LA, Jackson RD, Bartels RL, Lehnhard HJ. Weight-training effects on bone mineral density in early postmenopausal women. *J Bone Miner Res.* 1992;7:179–185.
26. Ryan AS, Nicklas BJ, Dennis KE. Aerobic exercise maintains regional bone mineral density during weight loss in postmenopausal women. *J Appl Physiol.* 1998;84:1305–1310.
27. Kelley GA, Kelley KS, Tran ZV. Exercise and bone mineral density at the femoral neck in postmenopausal women: a meta-analysis of controlled clinical trials using individual patient data. *Dis Management Clin Outcomes.* In press.
28. NIH Consensus Development Panel on Osteoporosis Prevention Diagnosis, and Therapy. Osteoporosis prevention, diagnosis, and therapy. *JAMA.* 2001;285:785–795.
29. Beck BR, Shoemaker MR. Osteoporosis: understanding key risk factors and therapeutic options. *Physician Sports Med.* 2000;28:69–84.
30. National Osteoporosis Foundation. *1996 and 2015 Osteoporosis Prevalence Figures: State by State Report.* Washington, DC: National Osteoporosis Foundation; 1997.

Received January 8, 2002
Accepted April 2, 2002

APPENDIX D

[AU: Please confirm MD degree]]

Exercise and Bone Mineral Density at the Femoral Neck in Postmenopausal Women: A Meta-Analysis of Controlled Clinical Trials Using Individual Patient Data

George A. Kelley, MD;[1] Kristi S. Kelley, MEd;[1] Zung Vu Tran, PhD[2]

The purpose of this study was to conduct a meta-analysis using individual patient data in order to examine the efficacy of exercise for improving bone mineral density (BMD) at the femoral neck in postmenopausal women. Ten controlled clinical trials that included 595 subjects aged 42–92 years met the criteria for inclusion. Across all designs and categories, there was an increase in BMD of 0.73%±5.52% and 0.45%±6.78%, respectively, in the exercise and control subjects. However, comparison of initial and final BMD values between exercise and control subjects revealed no statistically significant effect of exercise on femoral neck BMD. In addition, random-effects analyses revealed no statistically significant within- or between-group differences for studies in which individual patient data were available vs. those in which such data were unavailable. The results of this study suggest that exercise has no effect on femoral neck BMD in postmenopausal women.
(Disease Management and Clinical Outcomes. 2002;4:xx–xx). ©2002 *Le Jacq Communications, Inc.*

From the Clinical Investigation Program, MGH Institute of Health Professions, Boston, MA;[1] and the Department of Preventive Medicine and Biometrics, University of Colorado Health Sciences Center, Denver, CO[2]
Address for correspondence/reprint requests: George A. Kelley, FACSM, Associate Professor, Graduate Program in Clinical Investigation, Director, Meta-Analytic Research Group, MGH Institute of Health Professions, 101 Merrimac Street, Room 1059B, Boston, MA 02114
Phone: 617-724-5565
Fax: 617-726-8022
E-mail: gakelley@bics.bwh.harvard.edu
Manuscript received October 18, 2001; accepted November 9, 2001

Osteoporosis is a major public health problem in the United States. In 1996, it was estimated that approximately 29 million women and men over the age of 50 years had osteoporosis or had low bone mass and were at risk of developing the disease.[1] Of these estimated 29 million women and men, the vast majority (approximately 23 million) were women. By the year 2015, the prevalence and risk for this disease among women 50 years and older is estimated to increase to approximately 35 million.[1] The most devastating consequence of osteoporosis and osteopenia is an increased risk for fracture. For example, beginning at age 50, white women, a group at the greatest risk for osteoporosis and osteopenia, have a 40% chance of fracturing the spine, hip, or distal radius during their remaining lifetime.[2] In absolute terms, the number of fractures associated with osteoporosis has been estimated to be approximately 700,000 at the vertebrae, 300,000 at the hip, and 250,000 at the distal forearm. [2] In the United States, the health care costs associated with osteoporotic fractures exceed $13.8 billion annually.[3]

While the greatest number of fractures occur at the vertebrae, the most devastating fracture in terms of economic costs and mortality are fractures of the hip—specifically, the proximal femur.[4] For example, each hip fracture in the United States has been estimated to cost approximately $32,000 in total medical expenditures.[3] In addition, the 1-year mortality rate following a hip fracture is approximately 20%.[4] Furthermore, while difficult to assess, limited ambulatory skills, as well as fears about additional fractures, may affect the quality of life of individuals who have experienced a fracture.[1]

Finally, a woman's risk of a hip fracture is equivalent to her combined risk of developing breast, uterine, and ovarian cancer.[2]

A recent consensus statement from the National Institutes of Health Consensus Development Panel on Osteoporosis, Prevention, Diagnosis, and Therapy,[4] concluded that during the later years of life, exercise, in the presence of adequate calcium and vitamin D intake, has a modest effect on slowing the loss of bone mineral density (BMD).[1] Our previous meta-analytic research, using summary means from eligible studies, revealed improvement in BMD at the femur as a result of site-specific aerobic exercise, as well as progressive resistance exercise, in postmenopausal women.[5,6] Unfortunately, because of the small number of summary means from the studies, different assessment sites at the femur (femoral neck, Ward's triangle, trochanter, intertrochanter) had to be combined into one summary measure. Consequently, we did not examine the effects of exercise on BMD at the femoral neck, the most common site assessed at the femur. An alternative approach for dealing with this issue is to conduct a meta-analysis using individual patient data (IPD) rather than summary means from the studies. To the best of our knowledge, we are not aware of anyone who has attempted to use IPD for the purpose of conducting a meta-analysis on the effects of exercise on BMD at the femoral neck in postmenopausal women.

Given the health care consequences of low BMD at the femoral neck in postmenopausal women, as well as the potential benefit of exercise—a nonpharmacologic intervention that is available to most people—a need exists to examine the effects of exercise in this setting. Therefore, we conducted a meta-analysis using IPD in order to examine the efficacy of exercise in improving BMD at the femoral neck in this population of women.

METHODS

Data Sources

Studies for this investigation were extracted from a larger exercise and bone-density database in which IPD was available. Briefly, IPD were obtained by sending a cover letter and data request form to authors of reported studies via postal mail. Authors who did not respond received a follow-up request approximately 5 weeks after the initial mailing.

Study Selection

Inclusion criteria for this study were as follows: 1) randomized and nonrandomized trials that included a comparative non-exercise control group; 2) site-specific loading exercise lasting a minimum of 16 weeks; 3) postmenopausal women only; 4) journal articles, dissertations, and masters theses published in the English language literature; 5) studies published between January, 1966 and December, 1998; 6) BMD (relative value of bone mineral per measured bone area) assessed at the femoral neck; and 7) ability to obtain IPD from authors. Exclusion criteria for this study included the following: 1) studies in men or premenopausal women; 2) observational studies; 3) review articles; 4) case reports; 5) comments; 6) letters; 7) nonhuman studies; 8) foreign-language articles; and 9) abstracts from conference meetings. We did not include foreign language articles because of the potential for error in the translation and interpretation of findings. We limited our analysis to BMD at the femoral neck because this is the most common site assessed and because of missing data for the other sites (Ward's triangle, trochanter, intertrochanter). For studies that were ~~multiply~~ published, ~~i.e.,~~ as both a dissertation and a refereed journal article, we examined and referenced both in order to derive the maximum amount of information possible, but included this information as only one set of data.

Data Abstraction

A coding sheet that could hold 94 pieces of information was developed and used in this investigation. In addition, coding instructions that described how to code each item on the coding sheet were developed and used. The three major categories of variables that were coded for included 1) study characteristics; 2) subject characteristics; and 3) outcome data. All IPD were abstracted and checked for accuracy by the first author. Blinding of the coder to study information (identity and institutional affiliation of authors, study results) was not performed because it has recently been shown that these procedures have neither a statistically significant nor a clinically important effect on the results.[7]

STATISTICAL ANALYSIS

For IPD, means±standard deviations (SD) were used to describe continuous variables, while frequencies and percentages were used for categorical variables. For continuous variables, independent *t* tests were used to examine differences in initial characteristics between exercise and control groups, while chi-square tests were used for categorical variables. In order to examine initial and final values for BMD between ex-

ercise and control groups, a two-way analysis of variance (ANOVA) with repeated measures on one factor was conducted. The within-subjects or trial factor was assessment of BMD (initial and final), while the between-subjects or grouping factor was group assignment (exercise or control). Since this was an unbalanced design, a general linear model was employed. Prior to conducting the ANOVA, Pearson product moment correlations were used to examine for potential covariates to enter into the ANOVA model. This was accomplished by examining the association between changes in femoral neck BMD and the following variables: age, height, initial body weight, changes in body weight, initial body mass index (kg/m^2), changes in body mass index (kg/m^2), percent body fat, changes in percent body fat, initial lean body mass, changes in lean body mass, initial maximum oxygen consumption ($mL/kg^{-1}/min^{-1}$), changes in maximum oxygen consumption ($ml/kg^{-1}/min^{-1}$), age at menarche, years postmenopausal, cigarette smoking, alcohol consumption, calcium and vitamin D intake, compliance (percentage of exercise sessions attended), length of training in weeks, type of exercise intervention (weight-bearing, non-weight-bearing, weight training), type of BMD assessment (dual-energy x-ray absorptiometry, dual-photon absorptiometry), and study design (randomized vs. nonrandomized, controlled trial). However, since only initial BMD was significantly associated with changes in BMD (r=0.147; p= 0.011) and this variable was already in the ANOVA model, no covariates were entered into the analysis. The alpha level for statistical significance for both the Pearson product moment correlations and ANOVA was set at $p \leq 0.05$.

Because we were not able to retrieve IPD from all eligible studies, we also examined whether our results differed between studies in which IPD were available and those in which they were not. In order to accomplish this, we used the standardized difference effect size, calculated from the summary data reported in the studies.[8] This was calculated by subtracting the change outcome in the exercise group from the change outcome in the control group, and then dividing by the pooled standard deviation of the exercise and control groups. [8] The effect size was then corrected for small-sample bias.[8] If means and standard deviations were missing, the effect size was calculated from reported test statistics.[9] In general, an effect size of 0.20 is considered a small effect, 0.50 a moderate effect, and 0.80 a large effect.[10] An effect size of 0.50, for example, means that the exercise group differed from the control group by five tenths of a standard deviation in favor of the exercise group. For studies that included more than one group, an effect size was calculated for each group independently of the other. We used the standardized difference effect size vs. the original metric (BMD in g/cm^2) so that all eligible studies could be included in the analysis. We then compared effect size differences between those studies in which IPD were provided and those in which IPD were not provided, using an ANOVA-like random effects model developed for meta-analytic research.[8] This was accomplished by examining the between- (Q_b) and within-group (Q_w) differences for the effect sizes and their variances from each group. The alpha level for statistical significance was set at $p \leq 0.05$.

RESULTS

Study Characteristics

Twenty-three studies from 25 publications met our criteria for inclusion. [11–35] The number of publications exceeded the number of studies because we included two studies that appeared in the form of both a dissertation[29,33] and a refereed journal article.[30,34] From these, IPD was available from 10 studies and 11 publications.[12–16,21,22,27,29,30,32] Thus, we were able to include IPD from approximately 43% of eligible studies. A general description of the studies is shown in Table I. Five of the studies were randomized controlled trials and five were controlled trials. The 10 studies included a total of 22 groups (12 exercise, 10 control) and 595 subjects (295 exercise, 300 control) that met our criteria for inclusion. The subjects' ages ranged from 42–92 years in the exercise groups and 46–86 years in the control groups. The length of training for the studies ranged from 32–104 weeks (mean ± SD, 58±33 weeks). Ten groups performed weight-bearing exercise, two performed non-weight-bearing exercise, and two others performed weight training exercise. Six of the studies used dual-energy x-ray absorptiometry to assess BMD at the femoral neck while the other four used dual-photon absorptiometry.

Subject Characteristics

Initial characteristics of the subjects for continuous variables are shown in Table II. As can be seen, there were no statistically significant differences between exercise and control subjects in relation to age, height, body weight, and vitamin D intake. However, subjects in the exercise groups had significantly higher levels of calcium intake, while the number of years post-menopause was greater in the controls. Initial characteristics of the subjects for categorical variables are shown in Table III. No statistically significant differences were found between exercise and control groups in relation to cigarette smoking, alcohol consumption, use of estrogen

Table 1. Characteristics of Femoral Neck Bone Mineral Density Studies (gm/cm²) in Which IPD Were Provided or Available

Study	Design/Subjects	Exercise Intervention(s)	BMD Assessment
Bloomfield et al.[ ]	CT that included 17 postmenopausal women assigned to either an exercise (n=7; age=61.9±2.3 years) or a control (n=10; age=59.4±9.7 years) group	32 weeks of training performed 3x per week for 50 minutes per session (15-minute warm-up, 30 minutes of stationary cycling, 5-minute cool-down) at 60%–80% of MHR	DPA (Lunar DP3)
Bravo et al.[ ]	RCT that included 106 osteopenic women assigned to either an exercise (n=44; age=59.8±5.9 years) or a control (n=62; age=60.0±6.3 years) group	52 weeks of training performed 3x per week for 60–65 minutes per session. Exercise sessions consisted of a 10-minute warm-up, 25 minutes of rapid walking replaced with aerobic dance 1x per week, and 15 minutes of bench stepping at 60%–70% of MHR. This was followed by 10–15 minutes of resistance exercise.	DEXA (Lunar DPX)
Brooke-Wavell et al.[ ]	RCT that included 77 postmenopausal women assigned to either an exercise (n=8; age=64.9±3.0 years) or a control (n=39; age=64.1±3.1 years) group	52 weeks of training that consisted of self-monitored walking 3.5x times per week for 14.8 minutes per day for the first 12 weeks, followed by 20.4 minutes per day of walking, 4.8 days per week, for the remainder of the study.	DEXA (Lunar DPX)
Caplan & Ward[34]	CT that included 30 postmenopausal women assigned to either an exercise (n=19; age=66.4±5.1 years) or a control (n=11; age=65.4±5.0 years) group	104 weeks of aerobic weight-bearing exercise performed 2x per week for 60 minutes (warm-up, 20–25 minutes of low-impact aerobic exercise, 10 minutes of ball games for improved hand-eye coordination followed by work on floor mats for strength and flexibility, 10 minutes of relaxation). Subjects were also asked to exercise on their own 1x per week so that the pulse would be elevated for at least 20–30 minutes.	DPA (Lunar DPA)
Ebrahim[16] et al.	RCT that included 91 postmenopausal women assigned to either an exercise (n=46; age=66.5±8.1 years) or a control (n=45; age=68.2±7.9 years) group	104 weeks of walking 3x per week for 40 minutes per session.	DEXA (Lunar DPX)
Little[21]	CT that included 21 postmenopausal women assigned to a resistance training (n=6; age=59.3±2.4 years), walking (n=6; age=52.2±4.5 years), swimming (n=5; age=51.4±5.8 years), or control (n=4; age=60.5±1.3 years) group	Resistance training consisted of 32 weeks of training consisting of 9 exercises performed 3x times per week for 1 set of 8–12 repetitions at 60%–80% of 1RM; Walking consisted of 32 weeks of training, 3x per week for 30–50 minutes per session (5–10 minute warm-up; walking for 20–30 minutes; 5–10 minute cool-down) at 70%–90% of MHR; swimming consisted of 32 weeks of training, 3 x per week for 30–50 minutes per session (5–10 minute warm-up, walking for 20–30 minutes, 5–10 minute cool-down) at 70%–90% of MHR 13 beats per minute.	DPA (Lunar)
Lord[22] et al.	RCT that included 136 subjects assigned to either an exercise (n=66; age=70.9±5.0 years) or a control (n=70; age=71.0±5.0 years) group	42 weeks of exercise performed 2x per week for approximately 60 minutes per session (5 minute warm-up, 35–40 minutes of aerobic exercises consisting of activities for balance, hand-eye and foot-eye coordination, and strengthening exercises, 15 minutes of stretching, and 5–10 minute cool-down).	DEXA (Lunar DPX)
Prince[27] et al.	RCT that included 61 postmenopausal women assigned to either a calcium and exercise (n=26; age=63.6±4.5 years) or a calcium (n=35; age=62.4±4.8 years) group	104 weeks of weight-bearing exercise performed 2x week for approximately 60 minutes per session. Subjects were also asked to walk another 2 hours per week at 60% of peak heart rate for their age.	DEXA (QDR-1000)
Pruitt[29,30] et al.	CT that included 26 postmenopausal women assigned to either an exercise (n=17; age=53.6±4.1 years) or a control (n=9; age=55.6±2.9 years) group	36 weeks of strength training consisting of 13 exercises performed 3x week at 50%–60% of 1RM for 1 set of 10–12 repetitions for the upper body and 10–15 repetitions for the lower body.	DPA (Lunar DP3)
Ryan[33] et al.	CT that included 30 postmenopausal women assigned to either a weight loss (n=15; age=62.5±5.5 years) or an exercise + weight loss (n=15; age=63.4±5.7 years) group	24 weeks of aerobic exercise (treadmill jogging) performed 3x week for up to 35 minutes per session at 50% to >70% of Vo_{2max}. Each session included 10-minute warm-up and cool-down periods.	DEXA (Lunar)

Numbers of subjects are limited to those for whom valid IPD (individual patient data) were available. Groups were limited to those that met our inclusion criteria.

BMD=bone mineral density; CT=controlled trial; RCT=randomized controlled trial; DEXA=dual-energy x-ray absorptiometry; DPA=dual-photon absorptiometry; MHR=maximal heart rate; Vo_{2max}=maximal oxygen consumption; [AU: PLEASE SPELL OUT 1RM, and provide info for Lunar and QDR models: company, city, state.]

and/or progesterone, previous fractures at any site, and race. The majority of subjects were nonsmokers, not taking any type of estrogen and/or progesterone, and were white. A little more than one half of the subjects reported consuming alcohol and/or experiencing a previous fracture at one or more sites in the body.

Outcomes

Femoral neck BMD values are shown in Table IV. Across all designs and categories, there was an increase of 0.73%±5.52% in the exercise subjects and 0.45%±6.78% in the control subjects. Comparison of initial and final BMD values for potential differences is shown in Table V. Across all studies and subjects, statistically significant main effects differences were observed for groups but not trials. However, there was no statistically significant interaction between groups and trials, indicating that exercise did not have any effect on femoral neck BMD. In addition, no statistically significant interaction was observed when results were analyzed with each study and its subjects deleted from the model. Furthermore, no statistically significant within- or between-group differences in femoral neck BMD were observed when we compared effect size results for studies in which IPD were available with those that were not (mean±SD, IPD available, 0.098±0.352; IPD not available, 0.164±0.416; Q_w, 30.725, p=0.429; Q_b, 0.242, p=0.623).

DISCUSSION

The results of this study suggest that exercise does not improve femoral neck BMD in postmenopausal women. In addition, we found no differences between the results of those studies in which IPD were available and those in which IPD were not available. Furthermore, our results are consistent with the majority of findings of the original studies, in that 79% of the outcomes at the femoral neck were reported as not being statistically significant. In contrast, our findings appear to be somewhat different than the recent position statement from the National Institutes of Health Consensus Development Panel on Osteoporosis, Prevention, Diagnosis, and Therapy, which suggested that exercise during the later years probably has a modest effect on slowing the decline in BMD.[4] However, this was a broad statement and not specific to any one site in the body. Our findings also conflict with our previous meta-analytic work, in which an approximate 2% improvement in BMD was found at the hip as a result of site-specific aerobic exercise and progressive resistance training.[5,6] One of the possible reasons for the discrepant results between our current and previous meta-analytic work may have to do

Table II. Initial Characteristics of Subjects for Continuous Variables

Variable	N	Exercise Mean±SD	N	Control Mean±SD	Significance t (p) Value
Age (years)	278	64.90±7.11	291	65.13±7.12	−0.39 (0.70)
Height (cm)	277	159.01±6.69	288	158.42±7.08	1.02 (0.31)
Weight (kg)	278	65.54±11.28	286	64.77±12.26	0.77 (0.44)
Postmenopause (years)	133	14.37±9.97	127	18.60±10.67	−3.31 (0.001)*
Calcium (mg)	181	946.16±393.60	175	865.43±356.35	2.03 (0.04)*
Vitamin D (IU)	57	198±212	62	161±133	1.17 (0.24)

N=number of subjects for whom data were available; BMD=bone mineral density; t=independent t test value; p=probability value; IU=international units; *statistically significant

Table III. Initial Characteristics of Subjects for Categorical Variables

Variable	Exercise N (%)	Control N (%)	Chi-Square (p)
Cigarette smoking	19 (9.0)	25 (11.4)	0.71 (0.40)
Alcohol consumption	122 (59.5)	111 (52.9)	1.87 (0.17)
Estrogen/progesterone use	20 (6.9)	17 (5.9)	0.27 (0.60)
Previous fractures (any site)	58 (52.2)	63 (57.8)	0.68 (0.41)
Race (white)	213 (99.5)	214 (98.6)	0.98 (0.63)

Results are limited to studies that reported data for each variable.

with the fact that the summary measures obtained in our previous research were the result of pooling the outcomes from all sites assessed at the femur (femoral neck, Ward's triangle, trochanter, intertrochanter). Consequently, it may be that improvements in BMD occur at one or more sites other than the femoral neck. However, we were unable to address the effect of exercise on the other BMD sites at the femur in this investigation.

The recent consensus statement from the National Institutes of Health suggested that higher-impact activities and resistance training may have the greatest effect on BMD.[4] However, as can be seen in Table I, the majority of studies included in this investigation used lower- rather than higher-impact types of activities—primarily walking—as an intervention. In addition, while exercises designed to strengthen the hip were used in studies that employed a progressive resistance training protocol, the majority of exercises focused on movements designed to strengthen the upper body. Thus, it may be that the lack of improvement in femoral neck BMD in this investigation was the result of the exercise protocols employed. However, while higher-impact activities, such as jumping and high-impact aerobic dance, may be more beneficial to femoral neck BMD,[36,37] this has to be countered with issues of adherence to a regular program of exercise as well as the potential to put the subject at an increased risk for injury, particularly stress fractures-fractures and osteoarthritis.[38] Thus, from a practical standpoint, the lower-impact exercise protocols that were employed in many of our included studies are probably the most appropriate. This may be especially true for walking, since it is the most common type of exercise in which people in the United States participate.[39]

Since the terms lower- and higher-impact are broad and fairly subjective terms, it would appear plausible to suggest that future studies examining the effects of exercise on BMD make some attempt to quantify the forces involved for the interventions employed. For those studies that employ a progressive resistance training protocol, additional lower leg exercises that may affect femoral neck BMD should be employed. Incorporation of the above suggestions should result in a better understanding regarding the efficacy of exercise for improving BMD at the femoral neck.

Despite the fact that exercise did not have any effect on femoral neck BMD, such activities should almost always be recommended. For example, while exercise may not improve femoral neck BMD, it may increase muscular strength and balance and improve postural stability, thus reducing the risk of falling and the subsequent fractures that can result from falling.[4,40] While it is important for future research to examine the efficacy and effectiveness of various exercise interventions on femoral neck BMD, it would appear reasonable to suggest that a need exists for increased research that addresses the effects of exercise for preventing osteoporotic fractures in the presence and/or absence of changes in BMD.

While use of the meta-analytic approach provides for a more objective evaluation of studies rel-

Table IV. Femoral Neck Bone Mineral Density Results

Group	N	Initial Mean±SD	Final Mean±SD	Difference Mean±SD
Exercise	295	0.787±0.123	0.791±0.124	0.004±0.039
Control	300	0.763±0.122	0.764±0.117	0.001±0.048

Table V. ANOVA Summary Table for Femoral Neck BMD (General Linear Model)

Source of Variation	DF	SS	MS	F Ratio	P Value
Group (exercise and control)	1	0.193	0.193	6.730	0.010*
Subjects (group)	593	16.991	0.0287	—	—
Trial (initial and final BMD)	1	0.00217	0.00217	2.277	0.132
Group×Trial	1	0.000816	0.000816	0.857	0.355
Residual	593	593	0.565	0.000952	—
Total	1189	17.752	0.0149	—	—

ANOVA=analysis of variance; BMD=bone mineral density; DF=degrees of freedom; SS=sum of squares; MS=mean square; *P*=probability value; *statistically significant

young women following high-impact exercise. *Osteoporos Int.* 1994;4:72–75.
37 Welsh L, Rutherford OM. Hip bone mineral density is improved by high-impact aerobic exercise in postmenopausal men and women over 50 years. *Eur J Appl Physiol.* 1996;74:511–517.
38 Turner CH. Exercise as a therapy for osteoporosis: the drunk and the street lamp revisited. *Bone.* 1998;23:83–85.
39 US Department of Health and Human Services. *Physical Activity and Health: a Report of the Surgeon General.* Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1996.
40 Cummings SR, Nevitt MC, Browner WS, et al. Risk factors for hip fracture in white women. *N Engl J Med.* 1995;332:767–773.
41 Riggs BL, Melton LJ. The worldwide problem of osteoporosis: insights afforded by epidemiology. *Bone.* 1995;17(suppl):505S–511S. [AU: ORIGINAL LIST INCLUDED TWO REFS UNDER NUMBER 40. PLEASE CITE THIS REF (41) IN THE TEXT OR DELETE IT FROM THE LIST.]